

# FUSIO

The Bentley Undergraduate  
Research Journal

Volume: I    Spring 2017    Issue: II



Integrative and multidisciplinary undergraduate research  
at the leading edge of the Arts and Sciences and Business  
*Published by the Honors Program at Bentley University*



**BENTLEY**  
UNIVERSITY

### **EDITOR**

AARON L. JACKSON – ECONOMICS

### **ASSOCIATE EDITORS**

OTGONTSETSEG ERHEMJAMTS – FINANCE

JEFF GULATI – GLOBAL STUDIES

CHRISTIAN RUBIO – MODERN LANGUAGES

DAVID SZYMANSKI – NATURAL AND APPLIED SCIENCES

### **EXECUTIVE BOARD**

MIRIAM BOERI – SOCIOLOGY

DEBORAH GREGORY – FINANCE

DANIELLE HARTIGAN – NATURAL AND APPLIED SCIENCES

FRED LEDLEY – CENTER FOR INTEGRATION OF SCIENCE AND INDUSTRY

RACHEL MADDEN '18 – ACTUARIAL SCIENCES

KAREN OSTERHELD – ACCOUNTING

EUTHEMIA STAVRULAKI – MANAGEMENT

KATIE VADAKIN '17 – FINANCE

MATEUSZ ZEGLÉN '17 – MATHEMATICAL SCIENCES

### **PRODUCTION AND EDITORIAL ASSISTANCE**

MELISSA RUSH, MATTHEW DIMOND

---

### **BUSINESS AND EDITORIAL CORRESPONDENCE:**

BENTLEY UNIVERSITY

HONORS PROGRAM

175 FOREST STREET

WALTHAM, MA 02452

FUSIO@BENTLEY.EDU

WWW.FUSIOJOURNAL.ORG

---

---

## VOLUME I    SPRING 2017    ISSUE II

---

### ARTICLES

BRIDGING THE GAP: SECURING THE INTERNET OF THINGS MATTHEW C. AHLMEYER	3
THE IMPLICATIONS OF RELIGIOUS FREEDOM RESTORATION LAWS AND THE EVOLUTION OF FREE EXERCISE PROTECTION IN THE UNITED STATES AMANDA PINE	34
MANAGING STIGMA: WOMEN DRUG USERS AND RECOVERY SERVICES NAYEONG LEE AND MIRIAM BOERI	65
YAY OR NEIGH? IS THE HORSE RACING MARKET EFFICIENT? JILLIAN RAIA	95
ACKNOWLEDGEMENTS	131

## Aims and Scope

Established following Bentley University's mission of creating impactful knowledge within and across business and the arts and sciences, *Fusio* is a multidisciplinary undergraduate journal committed to the dissemination of original, high-quality undergraduate research. The journal is published by Bentley University's Honors Program and edited by both students and faculty across disciplines. *Fusio* encourages submissions from undergraduate students, with an emphasis on articles that span both business and arts and sciences topics as well as multidisciplinary topics. The journal is currently open only to undergraduate students at Bentley, and will consider original research by students as well as student/faculty joint work. All submissions undergo a blind peer review process.

*Fusio* is published twice yearly, in the fall and spring. For subscription information, please contact the Honors Program at Bentley University, or the journal office at [fusio@bentley.edu](mailto:fusio@bentley.edu). Online open access to these articles may also be found at [www.fusiojournal.org](http://www.fusiojournal.org).

Copyright © 2017 by Bentley University. Printed by Quad Graphics. Permission is granted to reproduce articles published in this journal, so long as reproduced copies are used for non-profit, educational, or research purposes. No responsibility for the views expressed by the authors in *Fusio* is assumed by the editors, by the publisher, or by Bentley University.

## BRIDGING THE GAP: SECURING THE INTERNET OF THINGS

By Matthew C. Ahlmeyer\*

*The Internet of Things (IoT), the next evolution in Internet technology, is predicted to add tens of billions of devices to the Internet in the next few years. IoT devices have been proven to provide major insights, collect valuable data, and make daily operations more convenient. Yet, there are inherent risks with IoT devices. Among the associated risks and challenges that must be addressed, IoT security is at the top of the list. Despite security being a major issue, IoT as currently constructed and designed is widely insecure. Literature emphasizes this point as authors are frequently stating that IoT security is a critical area of focus. Authors and researchers are constantly pointing out that security is an issue, but fail to address the possible solutions. Meanwhile, information security experts and IT personnel are writing other papers that look to offer technical, detailed suggestions to resolve the security issues, but often go without widespread implementation. Additionally, current security frameworks have not been updated to address IoT. This research will examine why this security gap is present within the Internet of Things consisting of a case study that analyzes several interviews with industry experts who work with IoT and information security on a daily basis.*

Keywords: Internet of Things (IoT), Information Security, Information Technology, Information Systems.

### I. Introduction

This research will study the Internet of Things (IoT) and the security and privacy issue that the industry is currently facing. The discussion will center around the ongoing conversation within the IoT community which continually points out security flaws that exist in newly created IoT devices and the reluctance of device developers to implement proper security protocols. The purpose of this case study research identifies the causes of the security gap present in Internet of Thing devices and addresses some ways to mitigate those risks while looking to answer the question, ‘What explains the gap between security desired and security delivered in IoT?’. Parts of the following

---

\* Email: mcahl14@optonline.net. I would first like to thank my capstone advisor, Professor Alina M. Chircu, for her continuous support and guidance throughout the completion of this capstone. Her immense knowledge and time spent were invaluable as I was writing my capstone. Next, I would like to thank each of the interview subjects for their time, knowledge, and contributions to the capstone. Without them there would not be any research or insight as to why there is this security and privacy gap in the Internet of Things industry. Finally, it is also essential that I thank United Technologies for their generosity in further funding my research. Traveling to a conference to present my work in front of industry experts and having my work published in an academic journal is something I am truly appreciative of.

paper were published in the *Issues in Information Systems* academic journal article “Securing the Internet of Things: A Review” written by Matthew Ahlmeyer and Alina Chircu.

The Internet of Things has arrived and is the next evolution—and disruption—in Internet technology (Dijkman et al., 2015; Ebersold & Glass, 2015). IoT (IoT) first appeared in literature in 2005 and is defined as “a system of interrelated computing devices, mechanical and digital machines, objects, animals or people that are provided with unique identifiers and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction” (Rouse, 2014). IoT utilizes technologies such as radio-frequency identification (RFID) and sensors to connect “things” in the environment to the Internet. In this context, a thing can be, for example, somebody’s heart monitor, a fitness tracker, an appliance, an industrial machine, or a car—all of which can collect data about their performance or location, save it and process it locally or on a server, and create alerts based on pre-defined rules—such as a car alerting the user when the tire pressure is low. Current projections state that the number of such IoT devices by 2020 will be over 26 billion (Lee & Lee, 2015). Part of the reason for this huge uptick in connected devices is the significant cost reduction in each of the components of IoT technology. Since 2010, sensor costs have been reduced by 50%, bandwidth costs have dropped by 40%, and processing power has been slashed by 60% (Hodgson, 2015).

However, with any opportunity there are also risks and challenges that must be addressed, and IoT security and privacy at the top of this list (Ebersold & Glass, 2015; Folk et al., 2015; French & Shim, 2016; Hodgson 2015; Weinberg et al., 2015). To provide context, it is important to note the difference between security and privacy. Security aims to provide protection for all types of information so that the confidentiality, integrity, and availability are all maintained. In contrast, privacy assures personal or corporate information is collected, used, and destroyed legally and fairly. Information technology (IT) security is already a major concern that businesses should be focused on—in fact, experts suggest that IT security considerations should be just as important as business considerations when making business decisions (Buecker et al., 2010). As businesses start implementing IoT devices to improve their internal processes and connect with their customers, IoT security becomes paramount (Accenture, 2015). However, while experts identify a significant need for IoT security, businesses in the IoT space, developers working for these companies, and users of their IoT solutions are either failing to implement security measures, or are not very concerned with the security issues (Weinberg et al., 2015). Often times this can be attributed to the fact that these businesses are newly established with little experience making these devices and are just looking to get their product to market as soon as possible. Thus, security features usually get looked over due to the added time and cost it would take to implement. Additionally, the convenience to not implement security features also plays a role as currently there is not any standard regulation for IoT devices, so developers opt to exclude these features. IoT adds a new dimension of security for businesses and users to

consider. Broadly stated, an unsecure connection could allow a hacker to not only access confidential information stored and transmitted through the device, but also could allow them access to other connected devices and data stored on the same network.

## II. Background

Recent Internet of Things devices are said to be physical, connected, and smart (Folk et al., 2015). These devices are said to be used by everyday people to manage their lives and business, connected to each other with the capability of sharing data, and are capable of analyzing data and information to enable autonomous actions (Folk et al., 2015). IBM has the viewpoint that there are two broad perspectives: Makers of Things and Operators of Things. Makers of Things focus on designing and manufacturing secure IoT systems and devices whereas Operators of Things utilize the created devices while still operating them in a secure manner. According to IBM, Makers of Things need to take a design for security and privacy approach meaning that devices need to be developed with end-to-end security built into the devices. This is a stark change from the personal computer era when the security culture was extrinsic or added-on after the fact through downloadable software across IT layers. IoT development teams need to spend the additional time to ensure sound security coding and features into IoT devices. This requires taking steps to harden devices. That includes changing default password to be complex, continually reviewing access control lists, and monitoring network traffic. Personally Identifiable Information (PII) needs to be properly handled by these device manufacturers to ensure that data cannot be directly linked to one specific individual. As for the Operators of Things, these users need to be aware of common security and privacy pitfalls and attempt to secure their network as best they can. This includes reviewing privacy policies and changing default passwords to limit the likelihood of their device—and ultimately their network—being breached. The risk of a breach is paramount because it creates the domino effect including decreased profitability, damaged brand reputation, and disrupted internal operations (Rothrock, 2015). Plus, a sophisticated attack will not stay within the company's walls, rather it can extend well beyond to other companies causing havoc at multiple levels (Rothrock, 2015).

The definition of IoT security is similar to that of mobile security which includes the protection of personal and business information that is stored, collected, and transmitted from devices connected to the Internet (Weber, 2015). This involves the protection from malware threats and unauthorized access to the device, in addition to facilitating attacks on other systems (Weber, 2015). Three major themes are present in the practitioner and academic papers investigated for our IoT security analysis. First, there are “call to action” papers which point out that security is an issue and those developing or utilizing IoT devices should pay careful attention to security. Second, there are technical papers which analyze IoT security threats and offer concrete solutions as to how to solve the security issue plaguing IoT. Third, there are legal frameworks papers

that look at the security and privacy regulations and laws and their application to IoT.

In recent years, cyber threats have grown exponentially in both quantity and volume (Hodgson, 2015). Security breaches and cyber heists are happening all around us and the authors of these papers do not expect that to change. This can and should be frightening to both companies and users. There are significant emerging security issues in IoT applications, networks, and devices/equipment, which could have major impacts on many industries and products (Accenture, 2015). Due to the unprecedented connectivity level in IoT, the potential vulnerabilities are also unprecedented (Folk et al., 2015), as every device is susceptible to being hacked and have its data compromised (Schneier, 2014). Hewlett Packard performed a study in 2014 that revealed 70% of IoT devices contain serious vulnerabilities (Lee & Lee, 2015). These security flaws are wide-ranging as “80% of devices failed to require strong passwords, 70% of devices did not encrypt communications, 60% lacked encryption for software updates and another 60% had insecure Web interfaces” (Britton, 2016). This was confirmed by Veracode who tested a handful of devices currently on the market as part of their IoT security research and found that all but one device tested was vulnerable in multiple categories (Veracode, 2015). Veracode tested a number of basic security features including the use of strong passwords, cryptography of communications, and certificate validation among others. Based on the results of the security tests, Veracode found that these devices are susceptible to account, network, and full service data breaches. With the growing concerns in cyber heists and the rapid growth of IoT, it is anticipated that 20% of security budgets will be designated solely to IoT by 2020 (Ranger, 2016).

A common phrase among practitioners is that security needs to be the DNA or the foundation of IoT (Folk et al., 2015). Companies involved in the IoT space can integrate security at the core of their value proposition by setting up a team of business executives and security specialists, integrating security best practice with IoT product development, educating customers and front-line staff in security best practice, and addressing privacy concerns with transparent policies (Turner, 2015). Four key areas for IoT security development need to be addressed: protecting communications, protecting devices, managing devices, and understanding your system (Symantec, 2016).

Alternative, more technical approaches, take aim at what security procedures and techniques should be implemented when developing these devices. First, a secure boot must be performed each time the device is turned on or activated. This is most likely done through proper cryptography methods. Next, proper authentication is essential through the use of strong passwords. Better yet, the use of X.509, an encryption authenticator, or Kerberos, another method of properly verifying the user, would provide a more secure authentication solution (IBM, 2015). Once the device and the user have been authenticated, secure communication must occur by the transmission of the data through secure encryption channels (SSH or SSL) (IBM, 2015). When done correctly, encryption can be extremely secure, however, there are many older forms of encryption that are less secure but popular to implement because of their simplicity. Protection against cyber-attacks and intrusion detection mechanisms must also be done

through the use of firewalls that limit communication to only known, trusted hosts (IBM, 2015). Additionally, embedding a device designed to detect and report invalid login attempts and other malicious activities (IBM, 2015). Last, but not least, the U.S. Federal Trade Commission (FTC) notes that basic, static security approaches cannot adequately secure an IoT device. It recommends that all devices be designed with continuous security procedures updates in mind, as security problems and solutions are always evolving (FTC, 2015).

However, consulting firms and security experts note that IoT devices inherently have some issues that would prevent them from performing some of these more technical security approaches detailed above. Aside from manufacture inexperience in dealing with security issues, IoT device developers may have a difficult—or perhaps impossible—time updating software vulnerabilities through patches. This could be due to the developer's inability to push the update to the device or the end users lack of attention to regularly checking for updates and patches.

### III. Methodology

This research examines the security gap currently present in the Internet of Things industry by interviewing industry experts. The basis of the methodology was derived from the book *Case Study Research Design and Methods* by Robert K. Yin. Yin developed a six stage plan for conducting research: plan, design, prepare, collect, analyze, and share.

Planning consisted of identifying a research question and identifying practitioner-generated and academic articles to review. Research and insight into IoT industry was conducted thorough interviews with industry experts who work with the technology on a daily basis. The interview subjects were carefully selected though personal contacts and referrals. These individuals have extensive business experience and are experts in information security—in particular they are familiar with Internet of Things devices. A broad range of subjects was preferred. This allowed each perspective of IoT security to be closely examined. For example, Chief Information Officers (CIO) or Chief Information Security Officers (CISO) were selected due to their ability to speak on how their business integrates IoT devices into their everyday business activities. The subjects were also selected from two different industries—financial services and education. This allowed for analysis to be performed between industries to see what similarities or differences each industry has regarding IoT security. Another interview subject focused more on the development stage of IoT devices. This individual works for Honeywell—a large manufacturing company that produces a wide variety of products including consumer products and engineering services for both private customers, major corporations, and the government. The questions focused more on asking the manufacturers at what stage in the process security is considered (if at all) and what features they are implementing (if any). The final cluster interviewed was security consultants from Accume Partners—an internal audit and risk advisory service firm—

located throughout the northeast. These individuals work in Accume's technology risk management and cybersecurity solutions practice. It is essential to interview these subjects because they are responsible for ensuring these devices are implemented correctly and safely to protect the data they collect.

A comprehensive literature review was performed using scholarly articles written by those in the industry—both individuals and company publications. Some of the articles were very technical in nature as they went into detail as to how IoT is currently operating and looked at the hardware, software, and network configurations at a very granular level. This was valuable for a basic understanding of the current set-up of IoT. Additional pieces of literature looked at the current state of IoT from a less technical lens. These publications were more concerned with stating that security is an issue, the reasons for it, and recent data breaches or vulnerabilities that exploited IoT devices. Aside from scholarly articles, five security frameworks were also reviewed to see how each handled securing and issuing controls to protect IoT devices. In all, 39 publications, articles, and frameworks were reviewed to understand IoT and the state of security.

Interview questions for the interviewees were developed using themes identified in the literature. The questions created were open-ended in nature to allow the interviewees to answer the questions freely. Additionally, the interviews were semi-structured. This means that the interview subjects were asked a set list of pre-prepared questions, but depending on their answers, the conversation could deviate in another direction based on their response.

**TABLE 1**

<b>Interview Questions</b>	
1.	What are the biggest benefits associated with the Internet of Things?
2.	What are the biggest risks associated with the Internet of Things?
2a.	What security flaws do you see relative to the Internet of Things?
2b.	Why do you feel there are potential vulnerabilities in IoT devices?
3.	What can be done to mitigate IoT device vulnerabilities?
4.	Are current security frameworks capable of monitoring and securing IoT devices?
4a.	If not: What requirements are necessary to include in an IoT framework?
4b.	Are you using COBIT, IBM, and other frameworks?
4b1.	How?
5.	Should there be government regulation on IoT devices?
5a.	Should there be industry standards?
6.	What IoT devices does your company utilize or develop? – provide examples (RFID tracking, sensors, etc.)
6a.	What practices are being taken to ensure data security?
7.	When developing/utilizing an IoT device what are the primary objectives/features of the device?
8.	How do you foresee the security issue being resolved?

9.	What techniques/strategies are you implementing in your business to mitigate IoT security risk?
10.	What costs are associated with IoT security? Worth the investment?
11.	Why do you think there is a security gap in IoT? Meaning security is an issue, but devices are still left vulnerable.
12.	Is there anything else you can think of that I haven't asked that would be helpful to my project?
13.	If something else comes up, can I contact you by email?

The collection of data was performed through a series of interviews conducted either over-the-phone or in-person. All except the Bentley University Chief Information Officer and Chief Information Security Officer were conducted via telephone calls. The date and time were mutually agreed upon by both the interviewer and interviewee a week prior to when the interview occurred. Each of the interviews lasted between 40 minutes and an hour. Interview subjects were made aware of the purpose of the research and gave permission to use their responses. Responses to the questions were documented during the interview using Microsoft Word. Finally, the interviewees also granted to be contacted by email should future questions arise later in the process.

TABLE 2

## Interview Subjects

Name	Industry	Company	Job Title	Interview Date, Time, Length	Interview Method
Michael Barrack	Risk Advisory Consulting	Accume Partners	Managing Director, Risk & Cybersecurity Services	11/30/2016 7:00 PM 45 Minutes	Telephone Conference Call with Robert Gains
Jason Christman	Manufacturing	Honeywell	Chief Cyber Engineer	11/8/2016 11:30 AM 1 Hour	Telephone Call
Robert Gains	Risk Advisory Consulting	Accume Partners	Senior Manager, Risk & Cybersecurity Services	11/30/2016 7:00 PM 45 Minutes	Telephone Conference Call with Michael Barrack
Colleen Medeiros	Financial Services	Bank Newport	Chief Information Officer	11/25/2016 9:30 AM 40 Minutes	Telephone Call
Ted Robitaille	Higher Education	Bentley University	Chief Information Security Officer	11/14/2016 10:00 AM 1 Hour	In-person Meeting with Bob Wittstein
Bob Wittstein	Higher Education	Bentley University	Chief Information Officer	11/14/2016 10:00 AM 1 Hour	In-person Meeting with Ted Robitaille

Following each interview the data and responses were reviewed and notes were clarified so they could be easily understood in the future. Data was analyzed following the completion of all interviews. Each interview was coded to identify themes both by question and for the entire interview. After looking for patterns, the data was sorted and grouped by theme for further analysis. Patterns were explored to see if these themes were consistent or inconsistent with the other interview subjects.

## IV. Results

Table 3 below provides a summary of the interviews conducted. This is intended to provide a high-level overview of the findings before a granular details and analysis is provided in each of the subsequent sub-sections. This table is also helpful in identifying themes discussed by each of the interview subjects along with similarities and differences between industries. Themes identified across industries or job functions include the human-side risk of IoT, lack of security framework application, and asset management.

TABLE 3

Interview Results						
Interview Subject Name(s)	Risks of IoT	Mitigate Vulnerabilities	Security Frameworks	Government Regulation / Industry Standards	Costs of IoT	Resolving Security Issue
<b>Michael Barrack &amp; Robert Gains</b>	<ul style="list-style-type: none"> <li>• People poor at security</li> <li>• Willingness to accept risks</li> <li>• Network connections and malware</li> <li>• Blind spot for companies</li> </ul>	<ul style="list-style-type: none"> <li>• Not accept risk</li> <li>• Asset management</li> <li>• Remote kill</li> <li>• Device registration</li> <li>• Firewalls &amp; event monitoring</li> </ul>	<ul style="list-style-type: none"> <li>• Not applied to IoT yet</li> <li>• Regulation vs application</li> <li>• Minimal level of security</li> <li>• Not designed to be technical</li> <li>• NIST &amp; FFIEC</li> </ul>	<ul style="list-style-type: none"> <li>• Regulation probably not necessary</li> <li>• Industry standards would be valuable</li> </ul>	<ul style="list-style-type: none"> <li>• Secure across many branches/sites</li> <li>• Growing complexity</li> <li>• High initial costs</li> </ul>	<ul style="list-style-type: none"> <li>• Education</li> <li>• Know what you are buying</li> <li>• Implement policies</li> <li>• Do not buy cheapest or take recommendations</li> </ul>
<b>Jason Christman</b>	<ul style="list-style-type: none"> <li>• Code developed by humans</li> <li>• Lack of testing</li> <li>• Only 1 vulnerability has to be exploited</li> <li>• Security perimeter</li> </ul>	<ul style="list-style-type: none"> <li>• Digital signatures</li> <li>• Rogue code</li> <li>• Stronger inputs into devices</li> <li>• Updates/patches</li> <li>• Risk score-cards</li> </ul>	<ul style="list-style-type: none"> <li>• Pull in appropriate frameworks</li> <li>• Use as a baseline</li> <li>• What best fits the company</li> </ul>	<ul style="list-style-type: none"> <li>• Regulation is a snapshot in time with limited foresight; stifles innovation</li> <li>• Industry standards always welcome</li> <li>• Regulate common processes</li> </ul>	<ul style="list-style-type: none"> <li>• Process of doing a risk analysis</li> <li>• Training developers and users</li> <li>• Customer support/manuals</li> <li>• Incident response</li> </ul>	<ul style="list-style-type: none"> <li>• Internet 2.0 revolution will resolve 90% of risks</li> </ul>

<b>Colleen Medeiros</b>	<ul style="list-style-type: none"> <li>• Devices simply lack security</li> <li>• Specific banking regulations</li> <li>• Middleware reliance</li> </ul>	<ul style="list-style-type: none"> <li>• Basic security techniques</li> <li>• Security reviews • Patching</li> <li>• Network layers</li> <li>• Vendor checklist</li> </ul>	<ul style="list-style-type: none"> <li>• Ineffective at addressing IoT</li> <li>• Look to SSAE16 and other audit reports, vulnerability assessments</li> </ul>	<ul style="list-style-type: none"> <li>• Regulation is too far</li> <li>• Business has to go on, but devices need security features included</li> <li>• Industry standards will probably come to financial services</li> </ul>	<ul style="list-style-type: none"> <li>• Security has to come first which takes time and money</li> <li>• Low return on investment</li> <li>• 70% will not get a return</li> </ul>	<ul style="list-style-type: none"> <li>• Waiting game until device are safer and more testing done</li> </ul>
<b>Ted Robitaille &amp; Bob Wittstein</b>	<ul style="list-style-type: none"> <li>• People, processes, software</li> <li>• DDoS</li> <li>• Design is not for security but usability</li> </ul>	<ul style="list-style-type: none"> <li>• Passwords</li> <li>• Asset management</li> <li>• Security reviews</li> <li>• Segment network</li> <li>• Veracode</li> </ul>	<ul style="list-style-type: none"> <li>• Ineffective at addressing IoT</li> <li>• Look to SSAE16 and other audit reports, vulnerability assessments</li> </ul>	<ul style="list-style-type: none"> <li>• Economics overrides all</li> <li>• Life safety issues</li> <li>• Higher education not very competitive already share info</li> </ul>	<ul style="list-style-type: none"> <li>• Convenience vs privacy</li> <li>• Hard to evaluate costs</li> <li>• Take offline if too vulnerable</li> <li>• Reputational risk</li> </ul>	<ul style="list-style-type: none"> <li>• Never will be fully resolved</li> <li>• Basic security techniques</li> <li>• Patches</li> <li>• Penetration testing</li> <li>• Education</li> </ul>

*INTERNET OF THINGS RISKS*

However, as has been seen in the personal computing revolution and the early stages of IoT, these devices are inherently unsafe. Each interview subject stated they could name numerous risks associated with both IoT as a computer architecture and the devices currently on the market. Jason noted one of the major root causes of risks in the IoT space is that software code is still developed by humans—something that will continue to occur for the foreseeable future. Due to human involvement in developing the software code for IoT devices, there will always be error, negligence, or ignorance during the development stage. While, testing will uncover some of the errors, there is not enough time or resources to find every bug in the code before the device is placed on the market. Additionally, even if 100 bugs are found in testing, a hacker only has to find one vulnerability or weakness in the code to exploit the device.

Ted Robitaille, Bentley University’s Chief Information Security Officer, made a similar point that these devices are often being manufactured to be attractive to the customer in terms of functionality or capability rather than their ability to be secure and robust. Robitaille pointed out that most of these companies manufacturing IoT devices are start-ups and are under serious time constraints from investors to get their product on the market. As a result, testing and security features are often left off to save money and time as these companies look to recover capital and become profitable.

Along the same lines, these security components are also removed to reduce the cost to the customer. As previously stated, it saves the manufacturer money by not investing in extensive testing, in turn, they can then pass along this reduced cost to the

customer. Most of the time, the customer does not see the value in security and is not willing to pay more for a device that is safer. Thus, they will often buy the cheaper of two products provided the functionality is the same. Companies also look to reduce cost by using cheaper inputs in their products. This includes using computer memories of minimal processing power or lesser storage capacity because as power or storage increase, costs also rise. Jason points out that this is not the wisest choice. Cheaper components often result in safety issues and are more prone to malfunctioning. This potentially could result in a hacker exploiting the device for a DDoS attack or the device failing at a faster rate due to inferior inputs. Among other malicious attacks, the interviewees expressed concern over race conditions, where one piece of code races and beats another piece of code triggering a series of glitches within the device, and replay attacks, where valid data is fraudulently repeatedly as part of a spoofing attack to flood the device until it ultimately fails.

Another risk of Internet of Things devices is the 3Vs property—volume, variety, and velocity—of big data. With the IoT and big data revolution, there is more data being collected than ever before. Also, this data is different than most other forms of data previously known. The information collected by IoT devices is a byproduct of so many different devices and formats which requires faster processing speeds due to the large quantities of data received and the importance of data collected. This has placed a burden on IoT device manufacturers and IT departments at companies as now there is more data to monitor. It is up to IT teams to be vigilant and take proactive steps to protect the data they are storing or the device they are manufacturing to protect the customer. As Jason said, this does not just include incident response, but also the architecture of the device and the network to protect the data from the beginning—not just after an incident occurs.

A common theme mentioned by the interview subjects was that with IoT a company's or individual's perimeter of devices or data no longer exists. Originally, a security perimeter could be viewed like a castle with a moat surrounding it. IT departments only had to be concerned about securing what was inside the moat as infrastructure and devices would all be operated and serviced in-house. The security perimeter then slowly started to disappear with the development of the laptop computer and mobile devices. However, IoT and cloud computing has completely removed the moat as there is no secure perimeter. Nowadays, the perimeter is dynamic and constantly changing, making it complex for IT departments to manage. This shift in security has resulted in device security becoming a focus of many IT executives. Firewalls and other perimeter security techniques are still absolutely necessary, but with so many entry points to the network, the end device has to be secure. Companies and individuals have devices connected to the Internet everywhere—including in their home and on their bodies—which make it more difficult to manage and keep track of them all. Also, if devices lack security, companies look to add middleware, but this further complicates the network and adds more vendors to monitor. IT departments need to be able to monitor all these devices to track data and watch for malicious attacks. However, as more devices get added to the network this task becomes more and more daunting.

Michael and Robert expanded on the number of connections to the network in their interview. While each noted some companies have elected to take the IoT and BYOD device route, they cautioned these organizations because their employees are using the devices are using them for more than just business purposes. Using them for personal use makes them more susceptible for breaches of company data especially if the user is connected to an unsecured home or public Wi-Fi connection. This can result in devices that become compromised by an authorized third party that can spread malware, download keystroke logging software, or take screenshots of what is occurring on the device. Companies try to encourage employees to use virtual private networks (VPNs) when connecting outside the office, however, they cannot easily enforce this. This creates a blind spot into what companies can and cannot control when it comes to data confidentiality, integrity, and availability. While the business can do all they can to protect the sensitive information and device when it is connected to the secure company network, they cannot govern what happens when the device is on a different network. This may leave the device vulnerable. The same can be said for IoT products. Depending on the network connection, the device may be left susceptible to hacks if the network is left unsecured or there are unencrypted communication channels between devices.

Finally, the last and ultimately biggest concern with IoT is with the end user. Michael and Robert explained that time and time again, some of the biggest security breaches are because of end users. As security consultants they find that people are notoriously poor at security. This includes writing down passwords on sticky notes because they cannot remember them, clicking on links in emails before checking that it was sent from an authorized source, or not performing updates timely. These are just a few common pitfalls of end users. With the implementation of more devices employees will be using, companies need to ensure that users are trained on proper security techniques and that they are practicing what is taught. It is ineffective to train employees if they continue to exhibit poor security practices. Companies also must conduct a cost benefit analysis to evaluate their comfort zone pertaining to their level of risk. If they find their employees and devices are at risk of exposing sensitive data, they should not venture into IoT products despite their potential business benefits. Finally, they must be committed to having a secure network environment with adequate security features in place.

#### *POTENTIAL WAYS TO MITIGATE RISKS*

With all of the threats IoT devices contain, it is important to examine what management and manufacturers are doing in attempt to mitigate risks. From a CIO and CISO perspective, both Bob and Ted emphasized the fact that it is extremely important for an IT management team to understand what devices they have. They stated that you can't secure what you don't know, so asset management is crucial when it comes to safeguarding devices. Both acknowledged that this process will continue to grow in

importance as companies continue to add devices to their network. Another step in the asset management process is to map the system and ensure that is constantly updated to reflect the most accurate representation of the IT environment. Michael and Robert agreed with Bob and Ted as they encourage their clients to register each device they purchase and add to their network. This way, they know exactly what they have to secure and how many devices are on their network. With more IoT devices, the map continually has more components, but having this understanding allows IT to comprehend the environment better which will allow them to address concerns easier should they arise.

A simplistic security approach, but one that is often overlooked with IoT devices, is using complex passwords when initially setting up devices. Bob and Ted mentioned that while passwords are weaker form of security, they still have potential in securing data and devices if used correctly. Where companies and people go wrong is that they fail to change the default password of the device. As a result, the default passwords are often very insecure and can be easily guessed by an unauthorized individual sometimes without the use of a password cracking algorithm. Employing best practice password controls with upper and lower case letters, numbers, and special characters that is over eight characters is the first step of configuring a secure device. This is even more paramount in IoT devices as due to their lack of computational power, passwords are often the only line of defense between the user and a hacker.

Another security technique deployed by Bob and Ted at Bentley University is to segment the network. This prevents a single point of failure from taking down the entire network. Utilizing proper segmentation or zoning can be effective in limiting further movement across the network should a security threat arise. Segmenting the network is done by creating firewalls and VLANs to deviate communication paths and set controls based on what information is traveling over the network. The security consultants agreed that setting up sophisticated firewalls and monitoring network traffic and events is essential. This way suspicious devices or activity can be picked up and stopped as soon as possible to prevent the malicious actions to continue. Bob, Ted, and Colleen also use the rule of least privilege when setting up devices on their company's network. This is when access is only granted to those who require the information for their job. Just because someone needs access to one area of the network, that does not mean they are granted full network access. Access is only provisioned to the areas of need and is continually reviewed and modified as roles change or people leave the company.

Along the same lines as segmenting the network is the ability to kill or wipe the device should it be lost or stolen. This means that the device will not permit anyone to log on to access the device's data or the data collected will be permanently deleted to prevent someone with unauthorized access to view it. Devices could also be wiped of all information after a certain number of incorrect username and password entries as that may be a sign of someone trying to gain unauthorized access. Often IoT devices are not contained within the business' walls, thus, they have a higher likelihood of be-

ing lost or stolen. Having this backdoor ability to kill devices is a must for companies as having access to one IoT device can open up the entire company network if not configured correctly.

Another area addressed by the CIOs and CISO was the management of third party devices and selecting a vendor. They all acknowledged that there are often multiple products on the market that can do virtually the same function, however, their organizations spend a lot of time choosing the vendor they are most comfortable working with. Security is one of the areas each of the executives looks closely at. All of them look for documentation of an external audit performed on the device when in discussions with the vendor. Typically they look to see that the device passed an SSAE 16 or SOC1. This proves that the device has been audited for minimum security levels and compliance with regulation. Bob and Ted also mentioned they will look to see if the device has been tested by Veracode for software code approval and if an ethical hack or penetration test has been conducted against the device prior to its release to market. Colleen also discussed BankNewport's checklist when selecting a device or application. The bank looks to ensure the device has basic security features—including complex passwords, timeout sessions, and biometrics—has an updated infrastructure, and is capable of patching the device should a bug be found, among others. In all, Colleen said the list has about 25 items, and the majority must pass in order for the device to be purchased. However, the bank will also perform walkthroughs of the manufacturer's facility if the device is considered to be high risk because financial data is heavily regulated and they have to ensure their customer's data will be safe. Deploying vendor management parameters and policies around mobile device administration is essential in ensuring that the vendors and devices selected are secure.

On the development side, Jason looks at how Honeywell goes about manufacturing the safest and most secure devices for their customers to use. Jason emphasized the importance of secure code or code that embeds an added level of security for the user within the device. Employees must be trained on how to securely code devices as this is the first step in creating a safe device. Because it all begins with device architecture, Jason said that security must be incorporated from the ground up. He also cited that extensive testing of devices must be done before selling the product as it only takes one vulnerability for a hacker to exploit. However, Jason stated that as an IoT device manufacturer becomes more mature and gets into a repeatable process of integrating security from the beginning, they may be able to skimp on testing because they have incorporated it from the outset and have experience in manufacturing safe devices.

Device manufacturers also have a heightened level of responsibility. Jason states that IoT device manufacturers should be providing detailed operation manuals to customers that explain how to properly install, configure, and operate the device so their data and information is protected as safely as possible. Additionally, manufacturers have to be responsible for the incident response of their devices. Jason explains that just because the product has left the manufacturer's office and is available for sale, it does not mean the manufacturer can forget about the device. In fact, Jason views this

as one of the most important stages of device development. Manufacturers must be pulling threat feeds, monitoring device traffic, and examining vulnerabilities long after the product has left their facility. If there is an issue with the device or there is unusual activity, the manufacturer must address the vulnerability and be able to push an update or a patch to the customer so they are no longer at risk. Often, this is the part where manufacturers fall short. They usually abandon their device after it has been sold and fail to provide updates even when a weakness is known.

Internal to the device, Jason discussed that the devices with the best security look to embed digital and cryptographic key signatures embedded within the device. This digital signature is unique to each user. If the digital signature is recognized by the operating system and processor, that proves the authenticity of the individual looking to access the device, and their information is processed or actions are permitted by the device. However, if the digital signatures do not align, the device restricts access as this is an unauthorized individual. This will prevent the device from operating rogue or malicious code. Devices need to be able to detect suspicious traffic and report that back to the user and manufacturer so both parties are aware of potential unauthorized access to their device. Finally, devices should also be capable of being patched if a vulnerability is found after its release. This would allow for updates to be pushed to the device to further mitigate a risk.

Jason also looked to provide some assistance for CIO and CISO-level management. He stated that a strong IT environment is where people, processes, and technology are all fused together operating as one cohesive unit. He stated this is accomplished by strong management who preach security from the beginning and always have it in mind when making business decisions. Additionally, Jason recommended the use of security frameworks and risk scorecards to identify threats when deploying IoT devices—something Bob, Ted, and Colleen are already doing at their respective organizations.

Finally, Michael and Robert noted that if a device is too risky or has too many vulnerabilities than it is not worth configuring on a company's network. During the vendor selection and cost benefit analysis performed in the initial stages of the process, a company will determine how much risk they are willing to accept with a given product. This varies depending on the function of the product, but if that level of risk is exceeded and the company is not comfortable accepting the added risk, they should abandon the project until a viable alternative is available within their risk level. While the function may be easier with the new technology, the ultimate concern of the company should be to keep their data and customers safe they should not take on the added risk.

### *SECURITY FRAMEWORKS*

When it came to the use of security frameworks, interviewees all agreed that the current security frameworks do not directly address IoT. That being said, their individual organizations are trying to use them the best they can. Starting with Bob and

Ted in the education industry, they both discussed how the current frameworks are very broad and have not yet been updated to include IoT technology controls. They continued to say that NIST is the best framework for addressing IoT, but it still has a little more to go before it will be more widely accepted. Bob and Ted added that they feel COBIT will also be modified to address IoT in the near future due to its wide acceptance in the information technology community. Bob and Ted say that in their jobs, they are personally using multiple frameworks and are applying them where they can depending on the framework's strength. They have found that CIS SANS Top 20 has been effective in providing security controls for the cyber defense strategy for Bentley.

Colleen echoed Bob and Ted's comments that current security frameworks have not yet addressed IoT. As a result, Colleen's team has had to utilize a mix-and-match approach when reviewing IoT devices. Colleen noted that BankNewport tries to limit adding IoT devices to their network due to their uncertainty, security vulnerabilities, and difficulties to monitor using security frameworks. She stated that banks are in a difficult position to utilize IoT devices because they are highly regulated by the government. As a result, they have to be careful what data they store due to its sensitivity and value to hackers. Nevertheless, like other industries, banks want to be able to better understand and serve their customers—something IoT devices enable—yet Colleen and other banks are hesitant to use some of this new technology due to government regulation and lack of framework guidance.

Colleen did state that the devices BankNewport utilizes throughout the organization are in the form of mobile applications. Similar to IoT devices, these apps have yet to be directly addressed in some frameworks. Colleen and her team must rely on the vendor to develop a secure app. Much like the vendor selection process, Colleen ensures the app manufacturer has passed and been issued an SSAE16 approval report. In addition, she reviews a checklist of minimum security features apps must have before the contract is signed.

Jason had similar comments as the management-level interviewees as he noted that organizations have been placed into a tough spot when trying to safely implement Internet of Things devices. Currently, he states that it is up to the company to pull in appropriate framework guidance and controls as a baseline and overlay them over the organization's IT environment the best way possible. This typically requires using multiple frameworks to best fit the company's situation as companies are using the most diverse set of technology than ever before. Jason accentuated that companies have a responsibility to analyze all existing frameworks to see what fits best along with ensuring that it is being utilized correctly to effectively protect the data they are collecting.

The security consultant perspective is especially valuable here as they are very familiar with the various frameworks and reference them on a daily basis. Michael and Robert state that there is currently an issue when it comes to IoT security regarding regulation versus application. As currently written, the security frameworks are a mere baseline of what should be done. They provide a minimal level of security and

do not provide a practical approach. The frameworks are effective at stating what to do, but are not technical in how to do it. Michael and Robert acknowledge that this is essential for the start when a business is new to the process of integrating security, but beyond that the frameworks do not serve a significant purpose. Many frameworks are taking a risk based analysis approach as opposed to trying to mitigate threats and best secure devices. Specific IoT features have not been included yet by the vast majority of frameworks which has left businesses to go with the mix and match approach Colleen discussed. With that being said, Michael and Robert agreed that NIST currently has the best approach to addressing IoT—especially considering its newest update released in November 2016—yet they still see areas for improvement. They also noted that the Federal Financial Institutions Examination Council (FFIEC) will probably be releasing something to help financial institutions govern their IoT cyberspace in the near future.

#### *GOVERNMENT REGULATION & INDUSTRY STANDARDS*

Without an adequate IoT security framework available for companies to use, the next question asked the interviewees if they felt government regulation would be beneficial to help secure IoT. Each of the interview subjects was conflicted on the question as there are positives and negatives associated with government regulation. Beginning with the pros, interviewees stated that regulation could result in an even playing field among all manufacturers keeping all of them accountable for developing devices that are safe and secure. Currently, manufactures lack an incentive to make devices secure because there is no regulation that is forcing them to meet a minimum standard or include a certain set of security features. Additionally, all subjects mentioned that if there is a national security issue or the safety of American citizens is at risk, then the devices should be regulated. Bob and Ted made comparisons of the FDA monitoring spoiled meat. The FDA monitors this because it is known that if humans consume spoiled meat there is a high likelihood of illness or death. The same applies to IoT. If the device puts human lives at risk, then it should be monitored and regulated, however, full regulation would be over the top for every device—especially if not life threatening. Michael and Robert also thought that there is not a huge need for regulation in the IoT space. They felt with proper implementation of IoT devices and general awareness of security practices that would be enough—assuming you have technical people working for the company who are capable of configuring the IT environment.

On the other side, there is a major negative should the government begin to regulate IoT devices. As has been proven in other industries, the moment regulation begins, innovation typically stifles. Ultimately, economics will override all. If it becomes more costly to include what the government is requiring the quality or uniqueness of devices will suffer as companies will spend more resources on compliance rather than novelty. Also, there are often adverse feelings towards being micromanaged. Both people and companies do not like when they feel they are constantly being watched every moment as then they lose their ability to work freely. Should regulations be passed, organiza-

tions would be burdened by the micromanagement of government officials. Clearly, there is a fine line that the government would have to balance so that business can continue to be profitable and innovative products can still be produced should regulation be enacted.

Additionally, Jason noted that government regulation is not always an effective means of regulation. This is because when governments pass laws they are put together by a bunch of regulators with little knowledge of IT. As a result, they receive guidance from IT advisors, but that might not be the best reflection of how the majority of professionals feel. Furthermore, regulation is a snapshot in time. By the time the laws are passed, there is a chance these regulations could already be outdated or will be outdated in the near future due to evolving technology or new threats. With this limited foresight, it would be necessary to change the laws frequently, however, this is both impractical and would not occur due to other concerns that must be addressed. Thus, the laws that would be passed would likely be broad and provide limited guidance to companies so they could be applicable to future technology. One area that has to be considered during regulation discussions is the issue of data localization. Data localization is storing user data in a data center on the Internet that is physically situated in the same country where data is collected. This creates issues for IoT because devices and sensors are located all over the globe. The question becomes who owns the rights to the data and where that data can be stored for future access. This is just one of the many areas of regulators would have to consider should they begin governing IoT.

Another point Jason raised that might be beneficial, but does not go as far as government regulation is looking at the common processes taken by manufacturing companies. Jason stated this is a viable alternative to government regulation because micromanaging the technological approach, stating exactly how the device needs to be configured, and what features need to be included will stifle innovation and restrict the novelty of devices. Instead, Jason feels there should be common processes each company has to go through when manufacturing a new IoT device. This includes considering and testing for security. That way this essential part of the process is not overlooked, but not so heavily regulated that innovation suffers.

Jason, Bob, and Ted are in favor of industry standards that would help govern IoT by industry. All three stated that each industry is different in their IT environment, strategy, and use of IoT devices, so rather than using a one-size-fits-all approach—like government regulation—industry standards might be more effective. Industry standards would allow similar companies to agree on best practices for manufacturing, securing, and using IoT devices in their industry. They agreed that the sharing of information can only benefit companies against vulnerabilities and hacks. In fact, Bob and Ted noted that they often have conversations with other CIOs or CISOs in the higher education industry to share information and tips. They also noted that this might not be as widely accepted in other industries because higher education is not as competitive and cut-throat as other industries. Yet, all three of these industry experts see benefits when working with others in the industry to develop a set of standards to safely protect organizations cyberspace and IoT devices.

As consultants, Michael and Robert see the benefits of industry standards, as well. Currently, Accume Partners serves a plethora of clients in the financial services industry—one of the more heavily regulated industries along with healthcare. As a result of serving so many clients in this industry, they have developed an expertise in this field. Having specific standards for each industry would allow for each company to compete on the same playing field and be accountable for the same set of standards. Additionally, it would allow for CIOs, CISOs, and consultants to specialize and focus on one set of standards for their industry rather than reviewing a dozen security frameworks to patch together a secure IT environment. In the United States there are currently some industry standards—including HIPAA for healthcare—but more widespread adoption of industry specific standards might be the most beneficial and practical option relative to information security as there seems to be a resistance to government regulation.

#### *COSTS OF THE INTERNET OF THINGS*

When including security in IoT devices, the additional cost that goes into including various security features becomes an issue. These costs include research and development, sophisticated or more complex hardware, and additional programming costs. The interviewees were questioned if the additional cost to manufacture or integrate an IoT device was worth it to their respective organization. Beginning with Jason on the manufacturing side, he felt as though security should always be one of the top priorities of companies manufacturing IoT devices—especially nowadays. Security determines how a product will age once it is exposed to the market. If the product is known for poor security and is exploited by hackers, the device's life cycle will be cut short and the manufacturer's reputation will suffer. Jason stated that manufacturers must perform risk analyses and ethical hacks to test how vulnerable their product is. If their product is vulnerable, they can either choose to upgrade the security features or risk that it may be hacked when it is connected to the Internet.

For Jason and Honeywell, the cost of spending the time to make their devices safe with adequate security features is not even questioned. Honeywell prides themselves on manufacturing quality IoT products. Plus, they have the capital and are willing to make the additional investment in their devices. It helps that they are an established company and can make this investment, but they also have people like Jason who are preaching security from the design architecture phase rather than rushing to get their product out the door. A common pitfall of many IoT device companies is that they are trying to recover start-up capital, so they forgo security to sell their product sooner.

From the management perspective, both Colleen from the financial services industry and Bob and Ted from the higher education industry had similar viewpoints on the costs of IoT security. Colleen was adamant that security is paramount and has to come first when looking at vendors and devices to implement. She would be willing to pay a little more if it meant she would be significantly more comfortable with the device's security. Bob and Ted agreed, but were not nearly as firm. Their different viewpoint is

likely because they are not handling highly sought after, confidential customer data. Each group also acknowledged the importance of keeping a strong brand image and reputation as it is embarrassing and damaging to suffer a breach. Additionally, they all stated that they perform a cost benefit analysis of each device they place on the Internet and they would be comfortable to take the device off-line if they feel it is too vulnerable. However, this is not always the easiest decision if the device is working effectively and is adding value to the organization. This constant battle with convenience and security is what these CIOs and CISOs are dealing with on a daily basis. Bob and Ted stated it is not the end of the world if the sprinkler system gets hacked, however, it is a concern if that can result in a larger breach of the network. They both stated that it is usually challenging to perform the cost benefit analysis and the cost of security is difficult to value. Frequently security is looked at like insurance in the sense that you are paying for something you never hope to use. Colleen added that 70% of the time you do not see the return on your investment when adding security features to your network or implement secure IoT devices. Bob and Ted agreed that often management—especially outside IT—do not see the benefits of security and only see the costs. Overall, both groups are willing to pay more for security, but for how much added protection is the question they are constantly evaluating.

When Michael and Robert are brought in to consult with a client, they see many of the same risk analyses stated by Colleen, Bob, and Ted. However, some of the other costs they see are the up-front costs of constructing a manageable yet secure network. It is an expensive process to purchase and configure the servers, routers, hubs, databases, cables, and other infrastructure components required to construct a network capable of handling IoT devices. The issue with adding IoT devices is that it adds a new level of complexity to the network and creates more possible entry points for unauthorized access. This is especially challenging for companies with multiple locations as the wireless network and communication channels have to be made secure across multiple branches or cities. There is often pressure to get it configured and finished as soon as possible, however, it must be done correctly and with proper security features or else there will be additional costs in the future should a breach occur.

#### *RESOLVING THE SECURITY ISSUE*

The final question posed to the interviewees inquired as to how the experts thought the security issue would be resolved, or if not completely resolved, how this security gap would be addressed. Of the interview subjects, Jason has the most positive outlook for the future. He stated that with Internet 1.0 the goal was simple: connect the world. That goal was accomplished, but the foresight was limited. During Internet 1.0, web pages were largely static and the current threat environment that the Internet possesses was not envisioned. Developers could not have dreamed of how many people would be using the Internet and they developed IPv4. There was virtually no consideration for security and there was a widespread adoption of poor security standards. This created

a problem during the initial stages of Internet 2.0 when the web became more interactive, social, and data driven. Security became more and more of an issue, but people were still using outdated standards in Internet 2.0. Everything was a bolt or an overlay, which has created numerous vulnerabilities. However, in recent years, there has been a large shift to improve security hence the development of IPv6. With the advent of Internet 2.0, it has become essential to account for devices on the network, identity and access management, and the storage and transmission of data. Jason feels the path that Internet 2.0 is taking will result in improved security. He estimates that 90% of the risks will be mitigated by Internet 2.0, but there will always be rebels or those looking to game the system, so the risk will never completely cease. That coupled with the general fear of suffering a cyber breach, adapting networks to handle mobile IoT devices, and including security during the initial stages of device development will result in a safer cyber space.

Michael and Robert see the security issue being resolved with two different education tactics. The first task of education is to have executive management and those in the information technology department be able to effectively train those using the technology and IoT devices. Individuals need to understand the proper ways to use IoT devices safely and securely. As stated earlier in the paper, people are inherently poor at information security and following secure practices even when mandated by their institution. Companies need to continue to invest in training and other programs for their employees. This will make them aware of the fact that they are constantly handling sensitive data and poor security practices jeopardizes the company. The second level of education is that companies need to know what they are buying. When a business is looking to implement an IoT device, they should be performing a detailed analysis of the device, including its benefits, risks, and features. Organizations cannot just look at the positives unless they are willing to accept the added risks. Going along with this, there should be technical people—mainly the CIO or CISO—present at these meetings when executives are looking at what devices to purchase. Security is no longer an IT issue, it is a company issue and needs to be discussed prior to the device being purchased. Those making key business decisions need to be educated on the device they are looking to purchase before making a final decision. Michael and Robert expanded to say that this includes not selecting the cheapest option or taking recommendations from another company. Often the cheaper products lack security features or would be impractical for the company. Even if it has great security, a \$50 router is not a viable option for a large company as that router is designed for consumer and home use not businesses. Additionally, Michael and Robert cautioned against taking recommendations. Other companies have different priorities, risk levels, and feelings about security, so what is acceptable for one company might not be the best solution for another.

Bob and Ted were convinced that the security issue will never fully be resolved, but see opportunities for it to improve. They feel that security will always be an issue because the hackers or ‘bad guys’ are always looking for ways to bypass the ‘good guys’ efforts to keep them out. The hackers will always be searching for vulnerabilities

as long as confidential, sought out information is stored on the Internet and there is a market for this information. Human ingenuity will always look to get around controls and security. Additionally, cyber space is innately human created—from development to exploitation—and as long as humans are involved, there will always be a risk. This leaves the ‘good guys’ on the defensive and forces them to respond to the hackers despite all the preventative measures implemented. Because IoT is so new, there are multiple areas for improvement. The first area where IoT can improve is to include basic security features in these products. As companies and people become more aware of the security in these devices, they will make an effort to select the ones that balance security with convenience. Manufacturers will begin to implement these features which will make it harder for unauthorized access. Additionally, manufacturers will also look to lengthen their device life cycle by applying monthly or emergency patches as needed to correct device vulnerabilities to keep those using the devices safe. Bob and Ted also noted that network security improved when ethical hackers or penetration testers attempted to breach a company’s network to see where the vulnerabilities were located. They envision that penetration testing will expand to IoT in order to keep the devices secure. Finally, as previously mentioned, Bob and Ted feel strongly about education when it comes to security. With this added awareness, security will also improve.

Finally, Colleen mentioned that the security issues will be resolved with further research and testing of devices to determine what device best fits the company’s risk environment while also serving a practical business purpose. She said that this was the case especially in her industry—financial services—due to heavy regulations and compliance levels. At this point, the industry is unsure of what devices they can implement and what data they are permitted to collect. As a result, there is a great deal of testing on devices while waiting for regulation to state what financial service companies can and cannot do. Colleen stated that it is important for companies to complete their due diligence on the devices they are looking to purchase and implement as they are ultimately the ones responsible for their customer’s data and security. Colleen also pointed out that the ability for the manufacturer to patch the device is a vulnerability is found is a critical feature and would mitigate a lot of security risks. Colleen, Bob and Ted mentioned that security will improve as organizations begin to develop policies surrounding the hardware and software companies look to implement. This includes defining how to evaluate vendors and device options. Formalizing this policy will result in more effective device selection and place an emphasis on security.

## **V. Recommendations**

After conducting and analyzing the interviews there are several recommendations that can be provided to help bridge the gap between IoT and the security of the devices. These recommendations address various areas, some of which are device manufacturers, end users, and security frameworks.

*INTERNET OF THINGS DEVICE MANUFACTURERS*

There are opportunities for manufacturers and companies implementing IoT devices to begin improving the security culture. Beginning with manufacturers, it is recommended that they invest and spend the additional capital and time into security of their devices. Security needs to be considered at the device design phase, and be embedded in the hardware. Nowadays companies are aware of security and are looking to implement not only convenient devices, but secure ones. Spending time to secure code devices and include multiple security features not only benefits the device but builds a strong reputation for the manufacturer. The IoT device itself needs to be designed using security principles. This covers the sensors that capture data, the data storage mechanism, and the micro-controller or actuator capable of controlling the device behavior, processing data and establishing a network connection. Additionally, the connections between the device and the network need to use secure technologies. Constructing a secure product will also increase the device's lifecycle. This comes at a cost of spending more time in research and development and using higher quality and costly inputs, but in the long run both the manufacturer and user benefit by creating a safer cyberspace.

Manufacturers also have to be responsible for the duration their product is in use to ensure a safe security environment. This includes monitoring devices for vulnerabilities and patching them as soon as a risk is found. A manufacturer's hands are not clean after they have released a device, rather their work is just beginning. They have to provide customer support for proper installation and incident management should something occur. It is recommended that companies start to be more transparent in their manufacturing process and they hire personnel specifically designed to troubleshoot or help with customer support and other technical issues related to their devices.

*INTERNET OF THINGS DEVICE USERS*

Companies implementing Internet of Things devices can also positively impact and change the current security culture. While all of the interview subjects in this analysis performed risk analyses of products they looked to implement, it is recommended that all companies perform detailed risk analyses when looking at purchasing and implementing IoT devices. Only products the company is comfortable with—based on their risk tolerance levels—will be added to the network. This process will ultimately make the network more secure. With this, only secure products will be purchased and implemented. In turn, this will then put pressure on companies who are manufacturing insecure products. They can either implement security features to remain competitive in the marketplace or continue to develop insecure products and risk that they will not be purchased.

Additionally, companies need to be aware of other security activities that can benefit them during IoT device implementation. It is recommended that company's use IoT device secure identifiers—such as serial numbers—that can be traced back to the

device manufacturer and uniquely identify devices as part of a device management program. Identifiers can be linked to permissions schemes that define how the device is allowed to interact with its environment, and these permissions can be changed dynamically—either by users or if the device is compromised during a security breach. Device management programs are essential as more and more connections to company networks are added, IT personnel must be able to monitor what each device is doing and if it should be on the network.

It is also recommended that companies implementing devices contract security experts to help them implement these devices and create a secure network. There are growing numbers of cyber and information technology consultants who can implement fundamental security measures to limit the vulnerability of a company's network. Companies can do their part to limit a breach or cyber-attack by implementing features including complex passwords, timeout sessions, and event monitoring among many others, but having a security consultant occasionally review the network and implement the devices is encouraged. Consultants have valuable experience and are extremely knowledgeable of best practices, not to mention having an independent third-party review the IT environment is valuable to catch potential vulnerabilities or errors.

#### *INTERNET OF THINGS SECURITY FRAMEWORKS*

Another major recommendation with respect to IoT is the issue surrounding the current state of security frameworks. All of the interview subjects noted that the existing security frameworks do not adequately address IoT—aside from NIST's very recent revision—and are not able to provide technical solutions. An integrated security framework that considers each stage of the IoT process—including development, implementation, and support—is necessary and sorely needed to improve the current security environment. First, pre-existing frameworks and IT governance needs to be updated to include specific IoT recommendations. For example, the COBIT standard (Wal et al., 2012) should incorporate principles specific to the use of IoT devices in and by enterprises – including principles for the governance of the hardware, network and software components of the IoT solutions for all COBIT processes (“Evaluate, Direct and Monitor,” “Align, Plan and Organize,” “Build, Acquire and Implement,” “Deliver, Service and Support,” and “Monitor, Evaluate and Assess”). By updating these frameworks, they will become modernized and applicable to current IoT technology.

Additionally, it is also recommended that IoT specific frameworks be developed that expand beyond NIST's recent update. Currently, organizations such as the IoT Security Foundation (IoTSF) and the Industrial Internet Consortium (IIC) are currently working on their own IoT security frameworks. Developing these frameworks exclusive to IoT will create the greatest value to resolve the security issue. Existing frameworks are trying to fit IoT into what they have already developed, however, IoT is a technology never seen before. As a result, it cannot be directly overlaid into an existing framework. These new frameworks will address how to configure a network

with hundreds or thousands of devices or sensors located around in numerous locations and how to secure the data secured on them and their communications. There are a lot of elements to consider when creating an IoT specific framework, but it is essential to thoroughly create a framework to better safeguard the devices and data stored on them from external hackers and create a safer cyber environment. It may be in some governing organization's best interests to share information and join together as they create a set of IoT standards to ensure they are sufficient, comprehensive, and there is widespread adoption of their framework.

#### *INTERNET OF THINGS GOVERNMENT REGULATION & INDUSTRY STANDARDS*

It is recommended that United States government not begin to regulate IoT in the near future. Currently, the technology is evolving at a rapid pace and pushing regulation on these companies would only stifle innovation. Additionally, regulation would not age effectively causing it to become outdated quickly. The technology will continue to evolve, but the law will be based on a set point in time which will only create more problems in the future. Not to mention, the innovation of products would likely suffer. As a result, it is strongly recommended that the government not pass any laws that begin to regulate IoT despite the current security flaws and vulnerabilities that the majority of these devices contain

That being said, an area that could be explored is the creation of industry standards. In the past related to other topics, industry standards have been better received by those in the industry—especially if it means the government does not get involved. Often companies would rather work with similar companies and self-regulate to create a more even playing field than have the government get involved. Usually government regulations are stricter and the companies have less influence in the decision making process. If industry standards were introduced to better securing IoT it would likely result in more information sharing between organizations on how to better integrate and configure devices to the network. Additionally, common processes could be regulated so every company in the industry has to perform the same steps. Industry standards would also be able to focus on areas specific to a particular industry rather than an overarching law that may be ambiguous. The ability to focus on certain areas and processes based on various quirks in each industry allows for a more customized and effective management approach. As previously stated, the higher education utilizes some of these industry standards and information sharing practices already and Bob and Ted expressed their satisfaction with how it works. Thus, it is recommended that other industries explore the possibility of developing standards specific to their industry.

#### **Internet of Things Security Education**

Successful implementation of all the previous recommendations is heavily dependent on the education of professionals who are currently working or will work with IoT. The past decade has seen the importance of strong information technology controls and policies grow exponentially and within the information systems community

there is recognition of the importance of security. Despite this analysis, security is not presented as a functional requirement during systems analysis and design training. If developers are not aware that they have to design for security, they expose the information systems they develop along with the associated data to significant security risks that are harder to mitigate after development has ended. Also, if users are not aware of security they could leave holes in their network, misuse devices, and potentially expose sensitive data. As previously recommended, designing for security is a best practice at all IoT solution levels—not just the application, but the hardware and network levels as well. Thus, it is recommended that educational programs in all areas relevant to IoT, such as information systems, computer science, and engineering, among others. These programs should focus on IoT security topics and provide in-depth coverage from each area's perspective on how to best secure a device from the outset. This education will provide awareness of the importance of IoT security from every angle—from development to implementation to auditing IoT devices. Education complements all the other recommendations. Without the knowledge or understanding for why to do something, people will be left confused and struggle to grasp the importance of IoT security.

## **VI. Conclusion**

Organizations and people are always obsessed with the latest trends and IoT is one of the most recent developments in the technology industry. In its short lifetime, IoT has proven to be very convenient, but also very risky. Current projections continue to state how beneficial IoT can be in our daily lives and references its unlimited potential. Companies are leveraging IoT to better understand their customers while personal users are trying to make their lives easier. However, the risks related to IoT can be frightening. The current digital age has led to valuable information being collected, processed, and stored over the Internet. As a result for all this valuable data, hacking has become commonplace and an unfortunate reality. Adding Internet of Things devices only creates more opportunities for unauthorized parties to access a company's network and create havoc. There are many linked processes and if security is not up to date and frequently examined, a business can suffer serious consequences.

This research looked to examine the reasons for the security gaps present throughout IoT. As previously stated, many academic papers have stated that IoT has limitless potential, but has residual risks—one of which is security. These authors simply state that IoT security is an issue and state reasons why including identity theft, data breaches, and hackings. The end of these articles is typically a call to action that states developers need to consider security when developing new products or that information technology professionals will be required to address upcoming security concerns. However, these articles typically offer no solutions or ways in which security can be improved. Another angle typically explored focuses on technical solutions to IoT security. Due to the fact that these papers are more specialized in nature, the managers and users concerned with IoT security may have difficulty in understanding how to actually

apply the solutions in their environment. Last, there are security frameworks that have yet to be adopted to help govern IoT. After discussing this security issue with many industry experts who serve in various roles—including management, development, and consulting—there are a couple major conclusions. The conclusions reached are that there is a lack of security in current IoT environment—both in the design phase and the implementation—and there is a lack of detailed, specific IoT guidelines in current IT security standards.

Relating back to the initial research question, there are a few reasons as to why there are significant gaps in the IoT industry. First, there is currently no incentive for manufacturers to create secure devices. There is very limited regulation over cyber activity and none relative to IoT. Manufacturers can create any device they want and put it on the market regardless of what security features are included—or not included. Without incentive to integrate security into their products, manufacturers are forgoing security as it is an additional cost. As a result, products with serious vulnerabilities are being sold and used on a daily basis. However, the manufacturers are not the only ones at fault. The second reason for this security gap is that companies are not performing risk analyses or review products for security before implementing them. Even if the product is convenient and makes processes easier, if it is outside the company's risk tolerance level and has the potential to lead to a breach or expose data, it should not be implemented. Yet, companies are purchasing these risky devices. This adds to the security issue as this signals to manufacturers that security is second-tier to convenience. This balance between security and convenience is the major reason for the security gap with respect to IoT.

Other contributing factors to the security gap include a lack of general education about security or ignoring the security discussion that are occurring. People are always the weakest link and this holds true with IoT. As long as people are involved, there will be an opportunity because people are infamously poor at security. Constant education is the only way this can be mitigated unless everything becomes autonomous and removes human interaction. Another reason why the gap exists is the lack of frameworks that address IoT to help organizations govern their information technology environments. These frameworks were especially helpful during the personal computer revolution and would also benefit IoT once developed. The sooner these frameworks are developed the sooner they will provide guidance to organizations on how best secure their networks with IoT devices connected to it. Additionally, the development of industry standards may also help rectify the security problem as it would keep companies in the same industry accountable for performing the same actions and manage common processes.

Due to this rapidly evolving technology, there are a few interesting areas for future research within IoT. First, an area of potential research would be to examine manufactured devices to see if there are more security features included on recently created devices compared to older ones. This would be interesting to see if manufacturers are spending more capital in research and development to include security features after

IoT devices have been the targets of hacks, been proven to be insecure, and experts have pointed out their vulnerabilities. Another area of research could be done to examine the various security frameworks and how they are being adapted to address IoT. Researching how NIST's recent update and various other frameworks mitigate risks would be an interesting analysis to perform. This would show which framework is most commonly implemented in companies and if one is proven to be more effective than the others. Plus, this could help identify areas that are either addressed effectively or ineffectively across various frameworks which may help identify why there are security gaps in IoT devices. These are just a couple areas of future research related to IoT and security, but there are many more that future researchers could examine.

This research sought to advance the understanding of the security gaps present in the Internet of Things while examining it through a variety of lenses—including technical risks and vulnerabilities, economic, and adoption barriers. Technology is dynamic. It will continue to change and adapt depending on the wants and needs of companies and the general public. Additionally, with the growing presence of the Internet in our daily lives, cyber challenges and network security are only going to grow in importance. There is a high demand for IT professionals to combat these risks. This research was able to identify the reasons for the security gaps present within IoT and provide some recommendations to start bridging the gap between security and convenience. Currently, there is a struggle between convenience and security as people want usable devices to make their lives easier, but security also has to be considered. Ultimately, joint action on the part of the IoT solution providers, businesses using IoT devices, consulting companies, regulators, and educational institutions will be needed to place security at the forefront of the IoT conversation and develop concrete security solutions for IoT.

## References

- Accenture (2015).** “Security Call to Action.” Available: [https://www.accenture.com/t20160122T014933\\_\\_w\\_/us-en/\\_acnmedia/Accenture/Conversion-Assets/Microsites/Documents22/Accenture-Security-Call-to-Action-IOT.pdf#zoom=50](https://www.accenture.com/t20160122T014933__w_/us-en/_acnmedia/Accenture/Conversion-Assets/Microsites/Documents22/Accenture-Security-Call-to-Action-IOT.pdf#zoom=50).
- Ahlmeyer, M. & Chircu, A. (2016).** “Securing the Internet of Things: A Review.” *Issues in Information Systems*. Available: [http://www.iacis.org/iis/2016/4\\_iis\\_2016\\_21-28.pdf](http://www.iacis.org/iis/2016/4_iis_2016_21-28.pdf)
- Ardiri, A. (2014).** “Is It Possible to Secure Micro-Controllers Used Within IoT?” *Evothings*. Available: <https://evothings.com/is-it-possible-to-secure-micro-controllers-used-within-iot/>.
- Britton, K. (2016).** “Handling Privacy and Security in the Internet of Things.” *Journal of Internet Law*.
- Cisco (2016).** “Securing the Internet of Things: A Proposed Framework.” Available: <http://www.cisco.com/c/en/us/about/security-center/secure-iot-proposed-framework.html>.
- Dijkman, R. M., Sprenkels, B., Peeters, T., & Janssen, A. (2015).** “Business Models for the Internet of Things.” *International Journal of Information Management*, 25(6), 672-78.
- Ebersold, K., & Glass, R. (2015).** “The Impact of Disruptive Technology: the Internet of Things.” *Issues in Information Systems*, 16(IV), 194-201.
- FTC (2015).** “Careful Connections.” Available: <https://www.ftc.gov/system/files/documents/plain-language/pdf0199-carefulconnections-buildingsecurityInternetofthings.pdf>.
- FTC (2015).** “Internet of Things: Privacy and Security in a Connected World.” Available: <https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-staff-report-november-2013-workshop-entitled-Internet-things-privacy/150127iotrpt.pdf>
- Folk, C., Hurley, D. C., Kaplow, W. K., & Payne, J. F. X. (2015).** “The Security Implications of the Internet of Things.” *AFCEA International Cyber Committee*. Available: <http://www.afcea.org/committees/cyber/documents/InternetofThings-FINAL.pdf>.
- French, A. M., & Shim, J. P. (2016).** “The Digital Revolution: Internet of Things, 5G, and Beyond.” *Communications of the Association for Information Systems*, 38(1).
- Goldstein, P. (2016).** “NIST Unveils Guidelines to Enhance Security of IoT Devices.” *FedTech*. Available: <http://www.fedtechmagazine.com/article/2016/11/nist-unveils-guidelines-enhance-security-iot-devices>
- Hodgson, K. (2015).** “The Internet of [Security] Things.” *SDM Magazine*. Available: <http://www.sdmmag.com/articles/91564-the-Internet-of-security-things>.

**Herrod, S. (2016).** “Connected Devices are Easily Hacked. Why Aren’t We Holding Manufacturers Accountable? *Recode*. Available: <http://www.recode.net/2016/10/24/13387188/connected-devices-hacking-Internet-manufacturers-security>

**IBM (2010).** “Introducing the IBM Security Framework and IBM Security Blueprint to Realize Business-Driven Security.” Available: [www.redbooks.ibm.com/redpapers/pdfs/redp4528.pdf](http://www.redbooks.ibm.com/redpapers/pdfs/redp4528.pdf).

**IBM (2015).** “IBM Point of View: Internet of Things Security.” Available: <http://public.dhe.ibm.com/common/ssi/ecm/ra/en/raw14382usen/RAW14382U-SEN.PDF>.

**IERC (2015).** “IoT Governance, Privacy and Security Issues.” *European Research Cluster on the Internet of Things*. Available: [http://www.Internet-of-things-research.eu/pdf/IERC\\_Position\\_Paper\\_IoT\\_Governance\\_Privacy\\_Security\\_Final.pdf](http://www.Internet-of-things-research.eu/pdf/IERC_Position_Paper_IoT_Governance_Privacy_Security_Final.pdf)

**IIC (2016).** “The Industrial Internet Consortium’s Approach to Securing Industrial Internet Systems.” Available: [http://www.iiconsortium.org/pdf/IIC\\_Approach\\_to\\_Securing\\_Industrial\\_Internet\\_Systems.pdf](http://www.iiconsortium.org/pdf/IIC_Approach_to_Securing_Industrial_Internet_Systems.pdf).

**IoTSEF (2016).** Press Release: “Internet of Things Security Foundation Drives Plan for the Supply Chain of Trust. Available: <https://iotsecurityfoundation.org/press-release-Internet-of-things-security-foundation-drives-plan-for-the-supply-chain-of-trust/>.

**ISACA (2012).** “Control Objectives for Information and Related Technologies (COBIT) 5”. Available: <http://www.isaca.org/COBIT/Pages/COBIT-5-Framework-product-page.aspx>

**Kovacs, E. (2016).** “IT, OT Collaboration Key to Securing Industrial Networks.” *SecurityWeek.com*. Available: <http://www.securityweek.com/collaboration-between-it-ot-teams-key-securing-industrial-networks>.

**Lee, I., & Lee, K. (2015).** “The Internet of Things (IoT): Applications, Investments, and Challenges for Enterprises. *Business Horizons*, 58(4), 431-40.

**McCoy, S., Everard, A., & Jones, B. M. (2015).** Foundations of Information Systems Course Content: A Comparison of Assigned Value by Faculty, Recruiters, and Students.” *Communications of the Association for Information Systems*, 36(1).

**Otto, G. (2016).** “New NIST Working Group Born Out of IoT Complexities.” *Fedscoop*. Available: <http://fedscoop.com/new-nist-working-group-born-out-of-iot-complexities>

**Palmer, D. (2016).** “The First Big Internet of Things Security Breach is Just Around the Corner.” *ZDNet*. Available: <http://www.zdnet.com/article/the-first-big-Internet-of-things-security-breach-is-just-around-the-corner/>

- Preimesbeger, C. (2016).** “Wide Disparity Between Consumer, IT Pro Views of IoT Security.” *EWeek*. Available: <http://www.eweek.com/security/wide-disparity-between-consumer-it-pro-view-of-iot-security.html>
- Ranger, S. (2016).** “Internet of Things: Finding a Way Out of the Security Nightmare.” *ZDNet*. Available: <http://www.zdnet.com/article/Internet-of-things-finding-a-way-out-of-the-security-nightmare>.
- Risen, T. (2016).** “The Privacy, Security Risks of the Internet of Things.” *U.S. News & World Report*. Available: <http://www.usnews.com/news/articles/2016-01-22/the-privacy-security-risks-of-the-Internet-of-things>.
- Rothrock, R. (2015).** “Why the Cybersecurity Domino Effect Matters.” *Forbes*. Available: <https://www.forbes.com/sites/frontline/2015/05/18/why-the-cybersecurity-domino-effect-matters/#59c793b75def>
- Rouse, M. (2014).** “Internet of Things (IoT).” *TechTarget*. Available: <http://Internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT>.
- Salisbury, W., Ferratt, T. W., & Wynn, D. (2015).** “Issues and Opinions: Assessing the Emphasis on Information Security in the Systems Analysis and Design Course.” *Communications of the Association for Information Systems*, 36(1).
- Schneier, B. (2014).** “The Internet of Things Is Wildly Insecure - and Often Unpatchable.” *Wired.com*. Available: <http://www.wired.com/2014/01/theres-no-good-way-to-patch-the-Internet-of-things-and-thats-a-huge-problem/>.
- Symantec (2015).** “An Internet of Things Reference Architecture.” Available: <https://www.symantec.com/content/dam/symantec/docs/white-papers/iot-security-reference-architecture-en.pdf>.
- Turner, M. (2015).** “How to Secure the Internet of Things.” *ComputerWeekly*. Available: <http://www.computerweekly.com/opinion/How-to-secure-the-Internet-of-things>.
- Veracode (2016).** “The Internet of Things: Security Research Study.” Available: <https://www.veracode.com/sites/default/files/Resources/Whitepapers/Internet-of-things-whitepaper.pdf>.
- Vermesan, O., Friess, P., Guillemin, P., Giaffreda, R., Grindvoll, H., Eisenhauer, M., Serrano, M., Moessner, K., Spirito, M., Blystad, L-C., Tragos, E. Z. (2015).** “Internet of Things Beyond the Hype: Research, Innovation and Deployment.” *European Research Cluster on the Internet of Things*. Available: [http://www.Internet-of-things-research.eu/pdf/Internet%20of%20Things%20beyond%20the%20Hype%20-%20Chapter%203%20-%20SRIA%20-%20IERC%202015\\_Cluster\\_%20eBook\\_978-87-93237-98-8\\_P\\_Web.pdf](http://www.Internet-of-things-research.eu/pdf/Internet%20of%20Things%20beyond%20the%20Hype%20-%20Chapter%203%20-%20SRIA%20-%20IERC%202015_Cluster_%20eBook_978-87-93237-98-8_P_Web.pdf).

**Wal, K. V., Lainhart, J., & Tessin, P. (2012).** “A COBIT 5 overview.” *ISACA* Webinar Program. Available: [www.isaca.org/cobit/documents/a-cobit-5-overview.pdf](http://www.isaca.org/cobit/documents/a-cobit-5-overview.pdf).

**Weber, R. H. (2015).** “Internet of Things: Privacy Issues Revisited.” *Computer Law & Security Review*, 31(5), 618-627.

**Weinberg, B. D., Milne, G. R., Andonova, Y. G. & Hajjat, F. M. (2015).** “Internet of Things: Convenience vs. Privacy and Secrecy.” *Business Horizons*, 58(6), 615-624.

**Xu, T., Wendt, J. B., & Potkonjak, M. (2014).** “Security of IoT Systems: Design Challenges and Opportunities.” *Proceedings of the 2014 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, San Jose CA USA, 417 – 423.

**Yin, R. K. (2009).** “Case Study Research Design and Methods.” Fourth Edition. *SAGE Publications, Inc.*

# THE IMPLICATIONS OF RELIGIOUS FREEDOM RESTORATION LAWS AND THE EVOLUTION OF FREE EXERCISE PROTECTION IN THE UNITED STATES

By Amanda Pine\*

*The 1990 Supreme Court case Employment Division v. Smith spurred the passage of the Religious Freedom Restoration Act of 1993, a federal law that requires courts to use the strictest form of judicial scrutiny when analyzing free exercise claims. There is disagreement among legal scholars as to whether the RFRA changed the face of free exercise law in the U.S. to a new, more stringent standard, or whether the law returned the courts' method of analysis to pre-Smith standards. The federal RFRA also spurred the passage of 21 state religious freedom acts. Both the federal and state laws have changed the face of free exercise analysis within the judicial system over the last decade. Additionally, the passage of these laws has raised concerns about the potential for religious freedom acts to be used for discriminatory purposes, particularly on the basis of gender and sexual orientation. This paper attempts to analyze the evolution of free exercise analysis both before and after Employment Division v. Smith and the RFRA, and to determine whether religious freedom laws can be used as a mechanism for discrimination. Based on an analysis of case law and legislation, it appears that the RFRA created a new standard of free exercise analysis, rather than returning it to pre-Smith standards. There also appears to be a clear potential for religious freedom laws to be used for discriminatory purposes.*

Keywords: Religious freedom, free exercise, civil rights, discrimination, equal protection.

## I. Introduction

In Detroit, Michigan, funeral home employee Aimee Stephens, previously Anthony Stephens, was fired from her position as a funeral director and embalmer at R.G. & G.R. Funeral Home. She held this position for six years, until expressing her intent to begin her transition from a man to a woman. As she transitioned, she expressed her desire to wear the women's funeral director uniform rather than the men's uniform, which she always had worn previously. The company owners asserted that allowing her to ignore her sex, which is a "God-given gift," would substantially impede upon their ability to conduct business in accordance with their sincerely-held religious beliefs

---

\* Email: amanda.pine28@gmail.com. I would like to acknowledge the support of my capstone advisor, Professor Franklyn Salimbene, who made this research project possible and helped me throughout the duration of the project.

(*EEOC v. R.G. & G.R. Harris Funeral Homes*, 2016, p. 16). The Equal Employment Opportunity Commission (EEOC) decided to represent the transgender employee, claiming that her termination violated Title VII of the Civil Rights Act of 1964 (Maatman and Karasik, 2016). However, in an interesting development, the Federal District Court of East Michigan struck down the application of Title VII to the funeral home, claiming that the provision did not pass the mandated strict scrutiny requirement for free exercise cases (*EEOC*, 2016). The court found that despite the fact that the funeral home had no religious affiliation, Title VII would substantially burden the religious beliefs of the owners, and as a result the law must pass the strictest form of judicial scrutiny.<sup>1</sup> The court ultimately found that Title VII was not narrowly tailored as applied to the funeral home, and therefore the business would not be subject to Title VII in this particular situation (*EEOC*, 2016).

This Michigan court decision is not the only recent case to attempt to balance the fundamental rights of free exercise of religion and equal treatment. In Kentucky, a local Fairness Ordinance, which banned discrimination on the basis of sexual orientation, was struck down as it applied to a closely-held, for-profit corporation called Hands on Originals. This company prints t-shirts and other promotional items for small organizations. The Gay and Lesbian Services Organization (GLSO), which plans the City of Lexington's Pride festivals every year, went to Hands on Originals for their Pride festival t-shirts. Despite seeming willing to provide the t-shirts initially, Hands on Originals allegedly went back on their agreement with GLSO upon finding out the message that the Pride festival promoted. The local Human Rights Commission ruled that the refusal of Hands on Originals to provide services to GLSO violated the discrimination provision of the Fairness Ordinance (*Hands on Originals, Inc. v. Lexington-Fayette Urban County Human Rights Commission*, 2015). However, the Kentucky Circuit Court overturned this ruling, finding that the Fairness Ordinance violated the company's right to free exercise of religion. The court found that, as applied to the facts of the case, the city did not have a compelling state interest in forcing Hands on Originals to provide GLSO with t-shirts, because GLSO could simply go elsewhere for t-shirts. Additionally, the court felt that the company was not denying services based on the claimants' sexual orientation, but rather because the owners disagreed with the message the shirts would be promoting. Ultimately the anti-discrimination law failed strict scrutiny in this case as well (*Hands on Originals, Inc.*, 2015).

These Michigan and Kentucky court decisions stem from the federal Religious Freedom Restoration Act of 1993, which was passed in reaction to the landmark court decision *Employment Division v. Smith*<sup>2</sup> (Luchenitser, 2015). This federal law and the

1. The strict scrutiny doctrine is a test that gives enhanced protection to fundamental constitutional rights. This standard of review requires a state to prove that a law promotes a compelling state interest and is narrowly tailored. Examples of constitutional rights that are sometimes afforded strict scrutiny protection include the First Amendment right to free speech and the 14th Amendment right to equal protection (Siegel, 2006).

2. *Employment Division v. Smith* was a Supreme Court decision about a controlled substance law in Oregon that banned the use of peyote. Two members of the Native American Church used this substance for religious purposes, and were subsequently fired from their jobs as a result. They did not receive unemployment benefits due to their use of an illegal substance, and thus filed suit, claiming that the controlled substances law violated their First Amendment rights. (*Employment Division v. Smith*, 1990)

resulting tidal wave of state legislation that is modeled after it require that any law that creates a substantial burden on free exercise of religion, whether or not the law is neutral or generally applicable, must promote a compelling state interest and be narrowly tailored by the least restrictive means (42 U.S.C. §§ 2000bb-2000bb-4). Some legal experts feared that these laws would allow for the use of free exercise claims to discriminate against individuals on the basis of gender and sexual orientation (Luchenitser, 2015). These worries had been largely unfounded until recently, when several state and federal courts began to strike down anti-discrimination provisions in favor of free exercise, particularly as applied to businesses. This newfound trend of free exercise claims winning out over discrimination claims has begun to set a potentially dangerous precedent within the judicial system. As laws passed with the purpose of ensuring equal protection are devalued in favor of religious freedom, an imbalance between fundamental liberties may become apparent. It is unclear how and if courts will draw the line between protecting the fundamental right of free exercise of religion, while at the same time ensuring that all citizens receive equal protection and fair treatment under the law.

## II. Background: The Religious Freedom Restoration Act

The Religious Freedom Restoration Act (RFRA) was passed by Congress in 1993 in response to the 1990 Supreme Court case *Employment Division v. Smith*. The Act was passed with largely bipartisan support and little controversy as a means of restoring a higher level of protection for free exercise in the United States (Seeger, 1997). The *Smith* decision, which determined that neutral and generally applicable laws create a burden on free exercise of religion so long as the law is reasonably related to a legitimate state interest, was perceived to have lowered the standard of judicial review traditionally utilized by the courts for free exercise claims. The legislative and executive branches, in reaction to this controversial decision, passed the RFRA in an attempt to void the *Smith* decision (Ryan, 1992).

There is a general consensus that the RFRA was implemented in reaction to the *Smith* decision (Lane-Steele, 2015). However, researchers have tried to understand why *Smith* triggered such an extreme reaction from legislators. The courts consider religious freedom to be a fundamental constitutional right in the United States (*Cantwell v. Connecticut*, 1940). Typically, courts analyze potential violations of fundamental rights under a heightened form of review known as strict judicial scrutiny (Siegel, 2006). However, this was not always the case when the Court examined free exercise claims. The ruling in *Employment Division v. Smith* determined that neutral and generally applicable laws that burden a religious practice or belief are subject to rational basis review, a lesser degree of judicial scrutiny (Hamilton, 2015). Before *Smith*, many claim that the standard of review for free exercise claims was strict scrutiny (Ryan, 1992). The landmark Supreme Court cases *Sherbert v. Verner* and *Wisconsin v. Yoder* initiated the use of the compelling interest test for cases that have to do with infringe-

ments upon religious practices. When *Smith*, in 1990, deviated from this standard that had been used for almost 30 years, there was a public outcry (Seeger, 1997). However, legal experts have argued that this public despair was the result of a misunderstanding of court precedent prior to the ruling in *Smith* (Ryan, 1992). In fact, most case law other than *Sherbert* and *Yoder* utilized a less stringent form of review for free exercise claims against a neutral law. Thus, *Smith* may not actually have been such a large deviation from the past. Rather, the public outcry was a result of misleading communication by the Court and deceptive lobbying by politicians who supported the Act (Hamilton, 2015).

Some legal experts and the writers of the *Smith* decision have claimed that even without the implementation of the RFRA, *Smith* would not have been likely to significantly change the result of free exercise litigation. Most free exercise cases prior to *Smith* failed in the courts even with the compelling interest test in place (Ryan, 1992). However, opponents of the *Smith* decision felt that the precedent would have diminished religious freedom within our country, and thus a reaction by the legislative branch was necessary (Seeger, 1997). Whether or not the RFRA was an overreaction to *Smith*, it was passed in 1993 with the clear intent of re-instituting the compelling interest test as the standard of review in free exercise cases. However, politicians and judges who oppose the law claim the RFRA went beyond simply re-establishing strict scrutiny as the standard. It is argued that the Act created what is known as super-strict scrutiny, which requires laws that burden religious practices to promote a compelling state interest, and also be the least restrictive means of achieving that interest.<sup>3</sup> The least restrictive means requirement was not used in cases such as *Sherbert and Yoder*. The RFRA therefore requires a stricter standard of review than was ever used previously (Hamilton, 2015).

The initial version of the RFRA prohibited any agency, department, or official of the United States government (both at the state and federal levels) from substantially burdening a person's exercise of religion. However, if the application of a burden furthers a compelling state interest, and is the least restrictive means of furthering this compelling interest, the burden will be considered permissible by the courts (42 U.S.C. §§ 2000bb-2000bb-4). The text of the law was unique in that it specifically mentioned the *Employment Division v. Smith* decision, stating that the court ruling "virtually eliminated the requirement that the government justify burdens on religious exercise imposed by laws neutral toward religion" (42 U.S. Code § 2000bb). The Act explicitly states that it attempts to restore the standard of review utilized in free exercise cases prior to *Smith*, naming the specific court decisions *Wisconsin v. Yoder* and *Sherbert v. Verner* (42 U.S. Code § 2000bb). The statute would therefore require any free exercise

---

3. The strict scrutiny test, as it was applied in free exercise cases prior to *Smith*, typically involved balancing the promotion of a compelling state with the burdens placed on religious exercise. Some versions of strict scrutiny prior to *Smith* utilized a narrow tailoring requirement (Siegel, 2006). Super strict scrutiny, however, requires that a state promote a compelling state interest using the least restrictive means. This creates a greater burden on the state than simply promoting a compelling interest, or ensuring that the promotion of this interest is narrowly tailored (Hamilton, 2015).

suit to which the government is a party to utilize the strict scrutiny standard of review, in which the state must meet the compelling state interest and narrow tailoring requirements.

The Supreme Court initially seemed to reject the application of the RFRA in its rulings. In the ambiguous court decision *City of Boerne v. Flores*, the court seemed to strike down the constitutionality of the RFRA, finding its application to state laws and local provisions outside of Congress's enforcement powers (Blatnik, 1998). The Court also disagreed with the law's specific reference to the First Amendment of the U.S. Constitution, because the Act appeared to be heightening the level of protection afforded by the First Amendment (*City of Boerne v. Flores*, 1997). However, because it was the application of the RFRA to the states with which Court seemed to have a problem with, the RFRA was revised to exclude state and local governments from its application. Additionally, the reference to the First Amendment was removed from the law and was instead replaced with a statutory definition of free exercise (42 USCS § 2000cc-5). It was unclear if the revised version of the RFRA would be upheld in court until the 2006 Supreme Court case *Gonzales v. O Centro Espirita Beneficente Uniao do Vegetal*. Since *Gonzales*, courts have regularly upheld the application of the federal RFRA (Lane-Steele, 2015). The law has even played a key role in landmark decisions such as *Burwell v. Hobby Lobby*.

The *Hobby Lobby* ruling spurred a new wave of academic discussion about the RFRA. The RFRA applies to any "person" whose religious rights are being substantially burdened. There is no explicit definition of a person in the RFRA, but the application of the law had never been applied to a business until *Hobby Lobby*. The ruling struck down the contraceptive mandate of the Affordable Care Act as it applies to closely-held for-profit organizations if the mandate violates the company's religious beliefs. There was much controversy over this decision, as some liberal politicians and proponents of the ACA argue that it opened the door for using the RFRA as a defense for denying certain employees benefits. Some argue that the RFRA is actually violating the First Amendment through this application, as it is imposing the religious rights of employers onto their employees (Luchenitser, 2015). One could even consider this application of the RFRA, whereby employers are imposing their religious beliefs onto employees, as a violation of the Establishment Clause of the First Amendment.<sup>4</sup> The implications of *Hobby Lobby* can be seen in the previously-mentioned court decisions in Michigan and Kentucky, where businesses have successfully made RFRA claims against anti-discrimination laws.

In recent years, many states have passed their own religious freedom restoration laws. In fact, as of the end of 2015, twenty-one states have passed such laws (State Religious Freedom Restoration Acts, 2015). The enactment of these state laws has created

---

4. The Establishment Clause states that "Congress shall make no law respecting an establishment of religion..." (U.S. Const. amend. I.). By endorsing certain religious beliefs or practices (such as the refusal to use or provide contraceptives) through the Religious Freedom Restoration Act, and imposing those beliefs or practices onto the employees of a non-religious company, it can be argued that Congress is violating this clause.

much debate and public outcry, with recent notable controversies in Arkansas and Indiana. Some claim that these laws create a mechanism for the discrimination of certain groups based on gender or sexual orientation, especially because some of the state laws apply to companies rather than just individuals (Hamilton, 2015). Additionally, some of the state religious freedom laws are not limited in their application to government laws and actions, but rather have opened the door to private citizen suits (Montanaro, 2015). Especially in places where there are not state-wide anti-discrimination laws for those who identify as a member of the LGBT community, the Religious Freedom Restoration Act creates a fear that people will discriminate and justify it with religious beliefs. This appears to be happening already, and it is unclear as to whether the legislature or the courts will do anything to remedy this newfound trend.

The controversy of the Religious Freedom Restoration Act, both at the federal and state level, ultimately comes down to a conflict between two important rights: freedom to exercise religion, as protected by the First Amendment, and equal treatment of individuals, as protected by anti-discrimination legislation and the 14th Amendment. Both the courts and legislatures have grappled with the issue of whether religious convictions can be used to justify behaviors that can be considered discriminatory.

The remainder of this paper will examine, using case law, legislation, and an analysis of today's political context, whether religious freedom restoration laws create a mechanism for discrimination. This paper will examine case law prior to the controversial *Smith* decision, as well as after the passage of the RFRA, to determine how the Court's analysis of free exercise claims has changed over the years. Additionally, this paper will examine regional and political trends regarding religious freedom restoration laws, and whether such laws appear to be in conflict with federal and local anti-discrimination statutes. Finally, I will recommend a mechanism for balancing the First Amendment protection of religious freedom with the fundamental right to equal protection and treatment under the law.

### III. Free Exercise Case Law

#### *BEFORE EMPLOYMENT DIVISION V. SMITH*

Even before the Supreme Court's controversial ruling in *Employment Division v. Smith*, the Court's analysis of free exercise claims changed substantially over the years. One of the first landmark free exercise cases was the 1940 case *Cantwell v. Connecticut*. This case dealt with a statute in the state of Connecticut which prohibited solicitation for money without the approval of a government official. The appellant in this case, Cantwell, was arrested after attempting to solicit funds in support of his religious order, the Jehovah's Witnesses. Cantwell claimed that this Connecticut statute violated his Fourteenth Amendment right to Equal Protection as it did not allow him to communicate his religious views. The Court ultimately found that the law violated both Cantwell's right to Equal Protection, as well as his First Amendment right to free exercise (*Cantwell*, 1940).

The major takeaway from *Cantwell*, however, is that despite a favorable ruling for the appellant, the Court explicitly discussed that not all aspects of free exercise rights are absolute. While the freedom to believe in whatever religion one chooses is absolute under the First Amendment, the freedom to act in accordance with those religious beliefs is not (*Cantwell*, 1940). The Court found that the statute, as applied to Cantwell, violated his religious beliefs because it was up to the discretion of one government individual as to whether he could solicit funds on behalf of his religion. In making this ruling, the Court also stated that “It is equally clear that a State may by general and non-discriminatory legislation regulate the times, the places, and the manner of soliciting upon its streets, and of holding meetings thereon; and may in other respects safeguard the peace, good order and comfort of the community, without unconstitutionally invading the liberties protected by the Fourteenth Amendment” (*Cantwell* at 304). Through this Fourteenth Amendment protection, the Court was also referring to a First Amendment protection of religious liberties.

Ultimately, *Cantwell* set the tone for the modern landscape of free exercise cases under the First Amendment. While free exercise is a fundamental right, religious conduct can be subject to reasonable, nondiscriminatory regulations by the state, particularly to safeguard the peace and good order of a community. There are certainly limits on the regulations that states can impose, as these laws must not unduly constrain religious conduct. However, as long as the state has reasonable interests and the regulations are neutral and not unduly burdensome, regulations on religious conduct are permissible (*Cantwell*, 1940).

This somewhat relaxed view on the regulation of religious freedoms changed fairly significantly with the Court’s 1963 decision in *Sherbert v. Verner*. In *Sherbert*, appellant, a member of the Seventh Day Adventist Church, was fired by her employer in South Carolina because she would not work on Saturdays, the day of the Sabbath for her religion. Appellant was denied unemployment benefits by the state of South Carolina because the state’s unemployment laws had a provision in which unemployment benefits are denied if the employee refuses available work. The Employment Security Commission found that because appellant refused to work on Saturdays, but would have been employed had she not refused to do so, she fell into the category of those declining available work. Appellant claimed that South Carolina’s Unemployment Compensation Act violated her First Amendment right to free exercise of religion (*Sherbert v. Verner*, 1963).

The Court ruled in favor of appellant, finding that the Act in its application to appellant violated her sincerely-held religious beliefs. The Court acknowledged the standards it set out in *Cantwell* by stating that “even when the action is in accord with one’s religious convictions, it is not totally free from legislative restrictions” (*Sherbert* at 403). However, for the first time in free exercise cases involving neutral and generally applicable laws, the Court implemented the use of the compelling interest test. The Court attempted to decide whether the state of South Carolina had a compelling interest to justify the enforcement of its unemployment benefits statute on the appellant, and whether such interests justify the infringement of appellant’s First Amendment rights (*Sherbert*, 1963).

The Court also explicitly rejected the use of rational basis review for the examination of free exercise claims.<sup>5</sup> “It is basic that no showing merely of a rational relationship to some colorable state interest would suffice” (*Sherbert* at 406). Despite moving away from rational basis review and implementing the use of the compelling interest test, the Court had not yet developed the narrow tailoring requirement that is present in many free exercise cases today. After *Sherbert*, the Court’s standard for free exercise cases was that the only permissible burdens on free exercise are ones that come from nondiscriminatory laws that incidentally burden religious practices in order to promote a compelling state interest (*Sherbert*, 1963).

The Supreme Court affirmed and further developed this standard in the 1972 case *Wisconsin v. Yoder*. This case was about the application of Wisconsin’s compulsory school attendance law to the Amish community. Respondents Jonas Yoder and his family were members of the Old Order Amish religion. In accordance with their religious beliefs, the family would not send their children to public school after the eighth grade, instead electing to provide the children with a private education that complemented their religious values. Respondents claimed that sending their children to high school of any kind would violate their sincerely-held religious beliefs, as doing so would expose the children to the world outside of their church community, endangering their salvation (*State of Wisconsin v. Yoder*, 1972).

In *Yoder*, the trial and circuit courts found that although respondents had sincerely-held religious beliefs, there was no violation of free exercise because the state of Wisconsin had a reasonable interest in requiring high school attendance. However, the Supreme Court of Wisconsin reversed, finding that the state’s interest in maintaining and establishing an educational system does not outweigh respondents’ rights to free exercise. Ultimately the United States Supreme Court upheld the decision of the Supreme Court of Wisconsin, finding that the Wisconsin attendance law unduly burdened respondents’ religious beliefs (*Yoder*, 1972).

The Court upheld the compelling interest standard initially utilized in *Sherbert*, when it stated that “only those interests of the highest order and those not otherwise served can overbalance legitimate claims to the free exercise of religion” (*Yoder* at 215). The Court also upheld the idea that regulations of general applicability can create undue burdens on religious conduct, as well as beliefs, which are both protected under the First Amendment. The Court essentially turned the compelling interest requirement from *Sherbert* into a balancing test of sorts, in which even if the state does have a compelling interest, as it did in *Yoder*, these interests must also outweigh any burden placed on religious conduct (*Yoder*, 1972). This balancing method was a further development by the Court in its approach to free exercise claims, and seems to be a step closer to the narrow tailoring requirement used today.

There appears to have been a clear trend throughout the 1960s and 70s of the Court heightening the protections it afforded to religious conduct, especially as it relates to

---

5. Rational basis review is the least restrictive form of judicial scrutiny, and only requires a state to show that a law is rationally related to a legitimate state interest to prove a law’s constitutionality (Harvard Law Review Association, 2012).

conflicts with neutral and generally applicable laws. However, in the 1988 case of *Lyng v. Northwest Indian Cemetery Protective Association*, the Court seemed to place a limitation on this increasingly heightened level of free exercise protection. While the Court upheld the idea that free exercise is protected under the First Amendment from both direct and indirect burdens, it also acknowledged that the government cannot account for every incidental burden that may be placed on any religion (*Lyng*, 1988).

*Lyng* was a case about the United States Forest Service's plan to create a 75-mile road in California. The National Forest Service (NFS) issued an environmental impact statement which found that the project would likely have an effect on the Native American groups who lived in the vicinity of the intended road plans. The NFS found that certain religious sites within the forest where the construction would take place could be affected by the project, thereby impeding the religious practices of these Native American groups. However, the Court noted that in response to the impact statement, those planning the project took actions to mitigate the impact the road construction would have on these groups. For example, protective zones were put into place around sacred sites. Despite this, several Native American organizations and individuals filed suit, claiming that the road project violated their free exercise rights because it would hinder their ability to practice their religious ceremonies in traditional ways (*Lyng*, 1988).

The Court found no violation of the Native Americans' free exercise rights. It re-emphasized that while religious conduct is certainly protected under the First Amendment, there are limits to this protection. The government cannot restrict religious conduct in any way, but it also cannot be forced to change its actions in a way that would conform to any particular religion. In a society as diverse as that of the United States, many government actions will invariably affect some religious conduct, and it is not possible to create accommodations for every single religious practice. The Court differentiated *Lyng* from its prior decisions in *Sherbert* and *Yoder*, as the statutes in those cases would have been in direct conflict with parties' religions. However, to rule against the government in *Lyng* would be forcing the government to relinquish property rights to conform to the religion of the Native Americans. The Court ruled that the government should not be forced to change its internal processes in order to accommodate a particular religion, thereby reaffirming a limit on the protections offered to religious conduct under the First Amendment. Because the government took steps to mitigate its impact on Native American sacred sites, and because the government cannot accommodate every single religion through its actions, the Court found no violation of First Amendment rights (*Lyng*, 1988). This tempered approach to the analysis of free exercise claims seemed to have reversed the trend of increasing protections for freedom of religion, and set the tone within the Court that influenced the *Employment Division v. Smith* decision a few years later.

*EMPLOYMENT DIVISION v. SMITH*

The Supreme Court's 1990 decision in *Employment Division v. Smith* was the Pandora's Box that ultimately resulted in the passage of religious freedom restoration laws. It changed the landscape of free exercise protections in the United States. *Smith* was a case about controlled substance and unemployment benefit laws in Oregon. Respondents were two members of the Native American Church who were fired from their jobs because they ingested peyote. Peyote is a controlled substance in Oregon under the state's criminal law statutes, but it is also a substance that is used for religious purposes by many members of the Native American Church. Upon being fired, respondents applied for unemployment benefits, which they were denied because they had been fired for employee misconduct. The respondents claimed that the controlled substance law and the denial of unemployment benefits violated their free exercise rights under the First Amendment. The Oregon Court of Appeals and the Oregon Supreme Court found that the denial of benefits did constitute an unconstitutional burden on free exercise. The Supreme Court reversed these rulings in an extremely controversial decision (*Smith*, 1990).

Justice Scalia, who wrote the majority opinion, stated that there was no violation of free exercise (*Smith*, 1990). Although legislators seemed to interpret the *Smith* decision to have drastically changed the analysis of free exercise cases, the Court seemed to believe that it was reaffirming its previous methodology for analyzing such cases. Rather than changing the standard through its decision, it was returning to the standard it had always used before the *Sherbert* and *Yoder* cases.

Justice Scalia asserted that the Court had never held that "an individual's religious beliefs excuse him from compliance with an otherwise valid law prohibiting conduct that the State is free to regulate. On the contrary, the record of more than a century of our free exercise jurisprudence contradicts that proposition." (*Smith* at 879). He then went on to cite a plethora of case law that supports the notion that religion does not exempt individuals from neutral and generally applicable laws. The majority opinion noted that to exempt individuals from neutral laws of general applicability would be to place religion above the law of the land, thus allowing any individual to become the architect of his own laws (*Smith*, 1990).

Next the majority opinion went on to differentiate previous cases in which First Amendment free exercise claims invalidated laws of general applicability from the situation in *Smith*. Justice Scalia asserted that in most of those cases, there were other constitutional protections at stake in addition to free exercise, such as freedom of speech, freedom of association, or freedom of the press. When cases involve not only free exercise, but other First Amendment protections, a stricter standard of review is warranted. For example, *Cantwell v. Connecticut* dealt not just with free exercise rights, but also with free speech rights related to Cantwell's ability to promote his religion through soliciting funds. Also, *Wisconsin v. Yoder* dealt with the rights of parents to raise children as they wish, in addition to freedom of religion (*Smith*, 1990). The

dual rights implicated within these cases warranted higher degrees of protection, and a higher standard of scrutiny. Consequently, the decision in *Smith* to use a lesser standard of review and to reject the free exercise claim of the respondents seemed to be consistent with the Court's historical decisions. Therefore, under this previously-unnamed hybrid rights doctrine, because the respondents in *Smith* did not present a secondary violation of any interest or right, strict scrutiny was not deemed the appropriate test (*Smith*, 1990).

The only missing puzzle piece that did not fit in with this hybrid rights standard was *Sherbert v. Verner*, which dealt only with unemployment benefits and a free exercise claim. However, the majority opinion states that the compelling interest test used in *Sherbert* has rarely been used outside of the context of unemployment benefits claims. Additionally, in the three occasions in which state unemployment benefit rules had been invalidated, those laws expressly conditioned the availability of benefits on "an applicant's willingness to work under conditions forbidden by his religion" (*Smith* at 883). At the time of the *Smith* decision, no law had ever been invalidated by the compelling interest test used in *Sherbert* outside of unemployment benefit laws. Additionally, the majority opinion noted that the Court had generally ceased using the *Sherbert* test outside of the context of unemployment benefits cases, specifically citing the example of the *Lyng v. Northwest Indian Cemetery Protective Association* decision (*Smith*, 1990).

The Court discussed the fact that even if it were inclined to begin using the *Sherbert* test again, it would never do so in a way that would invalidate a neutral and generally applicable criminal law, such as the controlled substances law in Oregon. The Court also created a loophole in which it could still utilize the compelling interest test for certain laws, particularly related to unemployment benefits. The Court stated that if a law sets up a system of particularized exemptions (as was the case with the law in *Sherbert*, in which an individual could still receive benefits if he had quit a job with good cause), then the compelling interest test is still applicable.<sup>6</sup> Other than in situations of hybrid rights violations, particularized exemptions, or certain unemployment benefits claims, the Court deemed the *Sherbert* test inconsistent with the majority of precedents. As a result, the compelling interest test was judged to be inappropriate for the analysis of free exercise claims that result from neutral and generally applicable laws (*Smith*, 1990).

While the decision was perhaps a bit complicated and nuanced, the main takeaway from the majority opinion in *Smith* is that the Court appeared to believe that its ruling was consistent with the majority of free exercise precedents. Legislators and lobbyists who denounced the Court's ruling in *Smith* portrayed the decision as a deviation from

---

6. In *Sherbert*, the Court determined that because someone could quit a job with "good cause," and still be eligible for unemployment benefits, this good cause standard constituted an individualized exemption to the statute. Ultimately, someone could refuse available work with good cause and still receive benefits. The Court stated that because an individualized exemption already existed in the statute's application, the exemption must be extended to apply to religious reasons for denying available work. Because the state had not applied the law in this way, strict scrutiny was the appropriate test. (*Smith*, 1990)

decades of free exercise precedent. Legislators viewed *Smith* as virtually eliminating constitutional protections for freedom of religion. It is clear from the majority opinion of the Court that this was not the case. While the compelling interest test had certainly been used in free exercise cases during the 1970s and 1980s in certain specific situations, it is clear based on the large quantity of case law cited by Justice Scalia that the use of this test was not the norm. Less stringent standards of review more similar to rational basis had typically been used to examine free exercise claims that stem from neutral, nondiscriminatory laws (*Smith*, 1990). As had been speculated by legal researchers in the past, it appears that the passage of the RFRA may have been due to a misunderstanding of the *Smith* decision, or perhaps an overreaction to it (Hamilton, 2015). Deceptive lobbying certainly could have played a role. If one were to believe the majority opinion in *Smith* and buy that most free exercise precedent did not support the use of strict scrutiny, then it is clear that the RFRA caused a significant deviation from precedent. This is in conflict with the RFRA mandating a return to stricter precedential standards, as many believed it was meant to do.

#### *AFTER EMPLOYMENT DIVISION V. SMITH*

Prior to the passage of the RFRA in 1993, the Court was able to affirm its decision in *Smith* in the 1993 case *Church of Lukumi Babalu Aye, Inc. v. City of Hialeah*. There, the Court upheld the tempered approach to the analysis of free exercise claims set up in *Smith*. The Court also showed that under the *Smith* standard, laws could still be found to unconstitutionally burden religion (*Church of Lukumi*, 1993).

*Church of Lukumi* was about the constitutionality of several ordinances passed in the City of Hialeah, Florida. These ordinances were passed after it was announced that a Santeria Church would be built in the city. This church uses animal sacrifice as a central part of its religious practice. After news of the church's establishment spread, the city held several town meetings to determine what course of action to take regarding the church's use of animal sacrifice. The city ultimately passed four ordinances that expressly banned the use of animal sacrifice in religious rituals. These ordinances specifically focused on animal sacrifice as it would occur as a part of the Santeria religion, but excluded all other types of animal killing. The Santeria Church filed suit, claiming that the restrictions on animal sacrifice violated their free exercise rights. The District Court and Appeals Court ruled in favor of the City of Hialeah, finding no violation of free exercise. The Supreme Court reversed these decisions (*Church of Lukumi*, 1993).

Even though the Court found that the ordinances violated free exercise rights, it utilized the standard set forth in *Employment Division v. Smith*. The Court discussed the fact that neutral, generally applicable laws are not subject to strict scrutiny for free exercise claims unless they are accompanied by some other legally protected interest. However, the Court in *Church of Lukumi* found that the ordinances in Hialeah were not neutral nor generally applicable and thus strict scrutiny applied. The ordinances in the city were passed to specifically target the Santeria religion, especially because they

banned religious animal sacrifice, but allowed all other types of animal killing. They also were passed with hostility toward the Santeria religion, as the town meetings prior to the establishment of the church made clear. It is apparent, based on these factors, that the ordinances were not neutral. The ordinances also were not generally applicable because they targeted a specific form of conduct that could harm public safety while neglecting similar forms of harmful conduct. The interests cited by the state, such as protecting children's emotional health, public safety, and preventing animal cruelty, would all require restrictions on other types of conduct besides religious sacrifice to truly be furthered. (*Church of Lukumi*, 1993).

The Court used heightened scrutiny to analyze the law because it was not neutral or generally applicable. *Church of Lukumi* demonstrated that the Supreme Court still took free exercise claims very seriously, despite the falsely perceived indifference toward free exercise after *Smith*. However, rather than seeing the positive aspects of these ordinances being struck down in favor of freedom of religion, some may have been more concerned over the fact that the Court had clearly adopted and utilized the standard set forth in *Smith*. Perhaps this is why the RFRA still passed with overwhelming bipartisan support later in 1993.

#### *Post RFRA*

Despite passing the Senate and the House of Representatives in 1993, the Religious Freedom Restoration Act did not receive much attention until the Court's 1997 decision in *City of Boerne v. Flores*. This case may have seemed on its face to be a typical free exercise suit; however, in reality it was one of the first major cases in which the respondent was making a claim under the RFRA, rather than under the First Amendment.

The Court focused very little on the facts of the case in *City of Boerne* and also very little on the free exercise claim. Rather, the Court conducted an in-depth analysis of the constitutionality of the RFRA. The Court found that the RFRA was an unconstitutional extension of Congress's enforcement powers under the Fourteenth Amendment, and that the provisions of the law that applied to the states were unconstitutional (*City of Boerne*, 1997). These findings aside, the decision left many wondering if any aspect of the RFRA was constitutional. The Court described the law as an inappropriate use of legislative authority, a violation of the separation of powers, and an attempt to change the level of protection offered to free exercise under the First Amendment (*City of Boerne*, 1997).

The Court did not feel that the RFRA was consistent with the protections provided by the First Amendment. "Legislation which alters the meaning of the Free Exercise Clause cannot be said to be enforcing the Clause. Congress does not enforce a constitutional right by changing what the right is. It has been given the power 'to enforce,' not the power to determine what constitutes a constitutional violation" (*City of Boerne* at 519). The Court seemed to interpret the mandate to use strict scrutiny for any free

exercise claim, even if it stems from a neutral and generally applicable law, as Congress's attempt to change the very meaning of the Free Exercise Clause. The majority opinion in *City of Boerne* described the RFRA as an attempt to make Congress's power paramount to that of the Constitution. The Court also discusses the fact that the need for the RFRA was nonexistent because, "RFRA's legislative record lacks examples of modern instances of generally applicable laws passed because of religious bigotry" (*City of Boerne* at 530). The Court's clear disapproval of the law, and of the changes Congress attempted to make to free exercise protections, made it unclear as to whether or not any part of the law remained valid.

The Court further discussed the dangerous precedent that the RFRA could set going forward. The RFRA mandates the use of the compelling interest test along with a least restrictive means requirement for the analysis of any free exercise claim. This is the most demanding test required of a state, and it is fairly difficult to pass. Many laws, the Court feared, would not meet this test. For example, the law in *Smith*, which banned the use of peyote, but was deemed constitutional due to its neutral and generally applicable nature, would likely have failed strict scrutiny under the RFRA. As a result, the RFRA could cause a flood of "constitutionally required religious exemptions from civic obligations of almost every conceivable kind" (*City of Boerne* at 534).

Finally, Justice Kennedy, writing for the majority, discussed the fact that the RFRA requires a level of constitutional protection for free exercise claims that was never offered before the *Smith* decision. As the Court previously discussed in *Smith*, most precedents did not use the compelling interest test for free exercise claims unless accompanied by the violation of some other legally protected interest. Additionally, even in the instances when the *Sherbert* balancing test was utilized, there was never a least restrictive means requirement. The Court interpreted the protections offered by the RFRA as broader than appropriate for the constitutional rights implicated, and as changing the very meaning of the protections afforded by the First Amendment. Justice Kennedy even went so far as to hint that because the legislative branch had passed a law that directly contradicted the "background of judicial interpretation," the Court would consider future free exercise cases in accordance with its precedent rather than in accordance with the new law (*City of Boerne* at 536). It was likely this harsh response by the Court to the RFRA that led many to believe that the law was declared wholly unconstitutional. It was not until nearly a decade later that the Supreme Court eventually upheld a revised version of the RFRA.

The Court's 2006 decision in *Gonzales v. O Centro Espirita Beneficiente Uniao Do Vegetal* brought the RFRA back to the forefront of the nation's attention. Since the Court's 1997 decision in *City of Boerne*, the RFRA had been amended significantly. The RFRA now applied only to federal laws, and legislators had cleverly removed the Act's reference to the First Amendment. Now, rather than expanding the powers protected by the First Amendment of the Constitution, the RFRA had its own statutory definition of free exercise (42 USCS § 2000cc-5). The mandated use of the compelling interest test could now be considered an additional protection for free exercise of reli-

gion, completely separate from those offered by the Constitution. Congress therefore attempted to preserve the separation of powers, while also heightening the required protection for religious freedom.

The Court's decision in *Gonzales* indicated that the Court noticed and agreed with the steps Congress took to revise the RFRA. *Gonzales* was about members of a religious sect called O Centro Espirita Beneficente Uniao do Vegetal who used a type of sacramental tea, known as hoasca, in religious ceremonies. This tea contained a substance that was banned under the Controlled Substances Act. When it was discovered that members of this religion were using the substance, they faced potential criminal prosecution. The group filed suit under the RFRA, claiming that the Controlled Substances Act substantially burdened their ability to practice their religion (*Gonzales*, 2006).

The Court ruled that the Controlled Substances Act violated the RFRA, and upheld the revised RFRA as constitutional. Until this case, the constitutionality of the revised RFRA had not been officially established. Additionally, in *Gonzales*, the Court officially rejected the framework for analyzing free exercise claims that was set up in *Employment Division v. Smith*. Instead, the Court used the mandated compelling interest test with the least restrictive means requirement to analyze the Controlled Substances Act. The majority opinion found that the U.S. government did not offer a compelling state interest to justify the burdens placed on the group's religious practices. The government had cited mitigating the health risks associated with using the tea, as well as compliance with an international treaty that regulates controlled substances as its two compelling interests. However, the religious group had maintained that it would only use the substance in controlled settings for religious purposes. The Court did not find these two interests compelling due to the fact that it saw little risk of health concerns from the group's careful use of the substance. The Court also felt that there almost no chance that the group's use of the substance would disrupt international relations, and so the second interest failed under the test as well. Finally, because exemptions had regularly been made to the Controlled Substances Act in the past (for substances such as peyote, which is used in religious ceremonies), an exemption could easily be made for the religious group in *Gonzales*. The Controlled Substances Act therefore failed the compelling interest test, and its application to the religious group's use of hoasca was deemed to violate the RFRA (*Gonzales*, 2006).

#### *BURWELL V. HOBBY LOBBY*

Since *Gonzales*, the courts have regularly heard cases under the RFRA. However, perhaps the most notable case in recent years was the 2014 Supreme Court decision *Burwell v. Hobby Lobby* which once again significantly changed the landscape of free exercise analysis in the courts. *Hobby Lobby* was a case about a provision of the Affordable Care Act that requires employers to provide health insurance plans to employees. These plans include coverage for women's health products which include some

types of contraceptives. Hobby Lobby Inc. and two other closely held, for-profit corporations filed claims under the RFRA, contending that the contraceptives provision of the Affordable Care Act substantially burdens their right to free exercise of religion (*Hobby Lobby*, 2014). Prior to *Hobby Lobby*, the RFRA had never been applied to a business before. The RFRA, according to the language of the legislation, applies to “persons” (42 U.S. Code § 2000bb–1). The law does not define persons, but Hobby Lobby claimed in this case that the term “person” includes corporations. In a controversial decision the Court agreed with the company and expanded the application of the RFRA to closely held, for-profit corporations (*Hobby Lobby*, 2014).

In *Hobby Lobby*, the Court again discussed the implications of the RFRA, admitting that the law provides a level of protection to free exercise rights that is greater than what is constitutionally required. Based on how broad this free exercise protection is, the Court found no reason to believe that Congress would have wanted to exclude small business owners from the law’s application. Because the law does not explicitly define the term person, the Court turned to the Dictionary Act for a definition. Under this Act the term person includes “corporations, companies, associations, firms, partnerships, societies, and joint stock companies, as well as individuals” (*Hobby Lobby* at 2768). Therefore, the Court felt that it made logical sense for the RFRA to apply to small, closely-held businesses, especially those that are owned by a single family with common religious beliefs. The implications of the Court’s ruling that “person” includes organizations could have significant consequences for the future, particularly if more and more businesses come forward with claims under the RFRA. It is clear that businesses have already begun to take advantage of the *Hobby Lobby* ruling, as the previously discussed cases in Kentucky and Michigan show. However, the Court also was very careful regarding the expansion of RFRA’s application, explicitly stating that the ruling in *Hobby Lobby* only applies to private, closely-held corporations. It did not expand the definition of person to publicly traded corporations or any other type of business. The Court would need to consider additional applications of the RFRA on a case-by-case basis (*Hobby Lobby*, 2014).

After determining that the RFRA applies to closely-held for-profit businesses, the Court analyzed Hobby Lobby’s free exercise claim under the compelling interest test mandated by the law. The Court did not question that Hobby Lobby’s owners had a sincerely held religious belief that life begins at conception. It concluded that forcing a company to provide health insurance coverage for forms of contraception that are contrary to this religious belief would substantially burden the owners’ free exercise rights. The Court initially began to question whether the contraceptive mandate is justified by a compelling state interest but then ultimately decided not to rule on that issue, choosing instead to focus on the least restrictive means requirement of the RFRA. The Court found that the Affordable Care Act fails the least restrictive means requirement because the Court could come up with alternatives to the framework set up under the Act that would not burden the religious beliefs of the business owners. These alternatives include exemptions that would exclude the contraceptive coverage from the

plans provided by Hobby Lobby and having the government pay for the contraceptive coverage for the women affected. Because the government failed to show why these exemptions are not possible, the Court found that the Act was not narrowly tailored and struck down its application to Hobby Lobby. The Court also justified this decision with the fact that there are already exemptions from the Affordable Care Act for non-profit, tax-exempt organizations and religious groups (*Hobby Lobby*, 2014). Overall, *Hobby Lobby* represents one of the most widely publicized and controversial claims under the RFRA to date.

The Court was extremely divided on the *Hobby Lobby* decision, and there were four justices who voted against the majority opinion. There were two full dissenting opinions written, one of which was an extremely vehement and detailed dissent by Justice Ginsburg. Her dissent seems to bring to light some of the major concerns that the public and certain justices on the Court had regarding the potential implications of the *Hobby Lobby* decision. Ginsburg stated, “In the Court’s view, the RFRA demands accommodation of a for-profit corporation’s religious beliefs no matter the impact that accommodation may have on third parties who do not share the corporation owners’ religious faith—in these cases, thousands of women employed by Hobby Lobby and Conestoga or dependents of persons those corporations employ” (*Hobby Lobby* at 2787) (Ginsburg, J., dissenting). Justice Ginsburg’s fear is one that has been reflected by the Court in previous free exercise decisions—that the creation of religious exemptions for any neutral, generally applicable law that incidentally burdens religion will cause religion to be held above the law of the land. It was concerning to the dissenting justices, as well as to the many citizens who opposed *Hobby Lobby*, that the Court does not consider the plight of the thousands of women employed by Hobby Lobby. These women will no longer have certain coverages under the Affordable Care Act as a result of the religious beliefs of a few business owners. These may even be religious beliefs that these women do not share. The *Hobby Lobby* decision thus creates a slippery slope regarding the power that religion may have over laws in the United States.

Justice Ginsburg’s dissent also expressed concern over the application of the RFRA to for-profit, commercial organizations. She argued that while the Court has recognized free exercise protections for non-profit, religious-based organizations in the past, it has never done so for commercial, for-profit organizations. She made a clear differentiation to show why free exercise exemptions should only apply to religious organizations. “Religious organizations exist to foster the interests of persons subscribing to the same religious faith. Not so of for-profit corporations. Workers who sustain the operations of those corporations commonly are not drawn from one religious community” (*Hobby Lobby* at 2798; Ginsburg, J., dissenting). Therefore, it does not make sense to dictate such organizations as if the stakeholders in these businesses operate under one religion. She also argued that had Congress intended to include businesses in the RFRA’s application, they would have explicitly done so. Additionally, the exemptions that are already in place for certain businesses under the Affordable Care Act only applied to religious non-profits, and not to commercial enterprises without religious

affiliations (*Hobby Lobby*, 2014). The fact that the Court was willing to equate exemptions created for religious groups with those created for non-religious ones could set dangerous precedent for the future.

Finally, Justice Ginsberg argued that the Court failed to even interpret whether the Affordable Care Act's contraceptive mandate *substantially* burdened the business owners' religious beliefs. While she did not doubt that their beliefs are sincerely held, she argued that the Court simply assumed that the burden created by the Act was substantial. Whether a burden is substantial is distinct from the issue of whether a belief is sincerely held. Justice Ginsberg argued that blindly providing health plans that include contraceptive coverage (with no guarantee that this coverage will be used, and no direct knowledge by the owners of the company as to who is in fact using this particular coverage) would not substantially burden the owners' religious conduct. She contended that any women who ultimately uses the coverage will not have been compelled to do so by the government or the business, and the decision will have been the woman's autonomous choice. The owners of Hobby Lobby would not have played a part in this autonomous decision, and therefore their religious convictions will not have been substantially burdened (*Hobby Lobby*, 2014; Ginsburg, J., dissenting). The *Hobby Lobby* decision therefore blurred the Court's previous definition of what constitutes a substantial burden under the RFRA, and this could again create a slippery slope for future free exercise claims.

Whether or not Justice Ginsburg and other dissenters' fears are warranted has yet to be seen. The far-reaching implications of the *Hobby Lobby* decision are just beginning to surface in the form cases such as the *Hands on Originals* and *EEOC* decisions discussed previously. Some of the most significant influences of the *Hobby Lobby* decision have been the influx of state legislation that mirrors the federal RFRA, as well as the national spotlight that has recently been placed on religious freedom laws. Furthermore, a widespread fear by lobbyists, legislators, and the public has surfaced regarding the issue of whether religious freedom restoration laws can be used as a mechanism to discriminate based on factors such as gender and sexual orientation. Because the Court has failed to draw a clear line about who can make RFRA claims, as well as how the law can be applied, much ambiguity still exists around the Act. As states continue to pass their own versions of the RFRA, the controversy around the issue of balancing free exercise with equal treatment of all members of society has heightened. This issue will only continue to gain public attention as more and more RFRA claims are heard in Court.

#### **IV. The Implications of State Religious Freedom Restoration Acts**

Because the Supreme Court ruled that the application of the RFRA to the states is unconstitutional in its 1997 *City of Boerne* decision, many states have passed their own version of the RFRA. In fact, since 1997, twenty-one states have passed their

own religious freedom restoration laws, some of which echo the federal law closely, and some of which deviate from the federal law significantly. Many of these state laws originate from the late 1990s, close to when the federal RFRA was deemed inapplicable to state laws. However, since the *Hobby Lobby* decision, the amount of proposed religious freedom legislation has increased substantially. In 2015 alone, seventeen states proposed religious freedom legislation to either create new laws or supplement already-existing ones (NCSL, 2015). Only two religious freedom laws were passed in 2015, in Arkansas and Indiana. Some of these proposed laws were vetoed as a result of widespread backlash, because many fear that these religious freedom laws will result in discrimination against minority groups under the guise of religious freedom (Montanaro, 2015).

While the Supreme Court has never ruled on cases under the RFRA that have resulted in discrimination, several lower courts have, as was the situation in the Michigan Funeral Home and the Kentucky Hands on Originals cases. Claims under religious freedom laws against anti-discrimination laws may increase—not only because the number of religious freedom restoration laws has increased, but because the number of anti-discrimination laws has increased in recent years as well. As the LGBT community has gained more rights and protections, the number of anti-discrimination laws that protect this group has grown. Particularly since the legalization of gay marriage in the 2015 Supreme Court case *Obergefell v. Hodges*, the increase in rights and protections for individuals on the basis of sexual orientation and gender identity has come to the forefront of national attention. While there are not yet federal laws that protect the LGBT community from discrimination, twenty-three states currently have anti-discrimination legislation that protects this group (ACLU, 2016). Even in states that do not have widespread LGBT protection, some cities and counties have their own ordinances that prevent discrimination. Interestingly, the influx in religious freedom laws and the increase in protections for the LGBT community have occurred at the same time. It is unclear if one of the trends has contributed to the other, or if they are related at all. However, the increase in both of these types of laws will likely create a challenge for both legislators and the courts as they attempt to balance the right to equal treatment with the right to free exercise.

One of the major issues regarding state religious freedom laws is the fact that they are often very different from the federal RFRA. Legislators, in writing and revising the federal RFRA, were very careful with how they constructed the law. They ensured that the religious protections provided by the RFRA were separate from those provided by the First Amendment. They also left the application of the law purposefully ambiguous in declining to define the term “person.” The text of the law made it clear that the test to be utilized for free exercise claims is the *Sherbert* test, with the added component of a least restrictive means requirement (42 U.S. Code § 2000bb–1). State religious freedom laws often deviate from these carefully-constructed standards. Occasionally, they provide a level of protection much greater than that provided by the federal RFRA. Some are also explicit in their application, greatly expanding the types of businesses

who can make claims under the state law. Some, in a rather extreme development, have even opened up the possibility for religious freedom acts to apply in private lawsuits between individuals (whether or not the government is a party to the infringement on free exercise). The apparent differences in state religious freedom acts and the trends that appear in this state legislation provide useful insight into the future challenges that may surround free exercise protection in the United States going forward.

## V. Trend Analysis: State Religious Freedom Restoration Acts

In the United States, 21 states have adopted religious freedom legislation similar to that of the federal RFRA. Interestingly, there is a fairly distinct regional trend regarding which states have adopted these laws. Of the nine states in the Northeast region of the country, only three, or one-third, have adopted religious freedom restoration laws. Similarly, only three states in the Midwest and three states in the West have adopted these laws. Meanwhile, in the South, eleven states out of seventeen states (including the District of Columbia) have adopted religious freedom laws. This is a large majority of states in a single region, particularly compared to the trend in other regions of the country. Ultimately, however, this tendency for Southern states to have religious freedom laws makes sense. States in the Southern region of the country tend to report higher levels of religious affiliation than states in other regions of the country (Lipka and Wormald, 2016). As a result, it is logical that the legislators in those states would make religious freedom a priority. While there is nothing inherently wrong with these highly religious states adopting legislation to protect their interests, a problem arises if the legislation adopted tends to favor a particular religion. In many Southern states, evangelical Protestants tend to dominate as the primary religious group (Wormald, 2015). If religious freedom laws favor the views of this particular group, then this could certainly create problems, especially if these laws are primarily used for claims against laws or government practices with which this faith tends to disagree. This section of this paper will attempt to analyze tendencies within state religious freedom laws to determine whether these laws are neutral to all religions, whether they create a mechanism for potentially discriminatory practices, and whether they inappropriately differ from the federal RFRA.

Many religious freedom laws do closely mirror the federal RFRA. While the federal RFRA is certainly not perfect and has been utilized in several controversial cases, the wording of the law is such that it would be hard to deny the law's neutrality. The RFRA was designed to add additional protections to free exercise of all religions. The wording of the law is purposefully vague, as is its application. The Supreme Court has continued to interpret the law in landmark cases such as *Gonzales* and *Hobby Lobby*, providing more clarity for the country on how the RFRA should be applied.

Many state religious freedom laws have maintained this neutrality and ambiguity of language, closely (if not exactly) mirroring the federal RFRA. For example, the Arkansas Religious Freedom Restoration Act is one that is very similar to the federal

RFRA. It explicitly mentions that the intent of the law is to restore the use of the compelling interest test utilized in Supreme Court cases such as *Sherbert v. Verner*. The law does not explicitly define to whom it applies. It uses the term “persons,” in describing its application, but much like the federal RFRA, it does not define what “persons” includes. The Arkansas RFRA also adds a least restrictive means component to the test it requires, much like the federal law does. Finally, the law does not draw its free exercise protection from the First Amendment, and instead uses a statutory definition to provide additional religious protection (2015 SB 975). This is important because the Supreme Court previously had a problem with the free exercise protections in these laws drawing their definitions from the U.S. Constitution. Ultimately, while the Arkansas law may face many of the same uncertainties and challenges as the federal law, it does not create further complications by significantly changing the language of the state law from the language of the federal law. In addition to Arkansas, states such as Alabama and Illinois also have religious freedom laws that are very similar to the federal RFRA.

Several state religious freedom laws are even more vague and ambiguous than the federal law is. Many of these laws do not mention the *Sherbert* test, do not define free exercise through a definition, and do not define the term “person” to clarify the law’s application. States such as Connecticut, Rhode Island, Arizona, Idaho, Kentucky, and New Mexico have laws with a bare minimum of language, and simply convey that they wish to heighten the degree of scrutiny for free exercise claims against state laws. Beyond that intention, these laws are vague enough to leave the majority of their interpretation up to the courts. Due to a lack of cases that utilize these laws, it is uncertain as to how courts will choose to interpret these vague and sparsely-worded religious freedom laws.

On the other side of this trend is the tendency for certain states to deviate substantially from the framework set forth in the federal RFRA. Several states have failed to adhere to the standard set forth by the Supreme Court in *Gonzales*, in which the Court only accepted the constitutionality of the RFRA because it created a protection for free exercise outside of that provided by the First Amendment (*Gonzales*, 2006). The federal RFRA used a statutory definition of free exercise to justify the strict scrutiny test. Similarly, state religious freedom acts that mirror the federal Act typically use either a statutory definition of free exercise or cite provisions within the state’s constitution that grant protections to religious freedom. Several state religious freedom laws do not follow this standard and, instead, draw their protection directly from the First Amendment to the United States Constitution.

For example, the Mississippi Religious Freedom Restoration Act defines free exercise as “the exercise of religion under the First Amendment to the Constitution” (Miss. Code §11-61-1). The problem with this definition is that the Court expressly stated in *City of Boerne v. Flores* that the Court’s precedent does not support the use of a compelling interest test with a least restrictive means requirement for free exercise claims that stem from neutral laws of general applicability. Based on decades of the Court’s interpretation of free exercise claims, the First Amendment does not warrant

this extremely heightened level of protection. The Court accepted the constitutionality of the RFRA once the reference to the First Amendment was removed. Based on this precedent, it seems as though, if the Court were to analyze the constitutionality of the Mississippi RFRA, it would not consider the Mississippi law to be constitutional. The Court did not seem to believe that Congress had the authority to create additional protections under the Constitution, effectively amending the document. It seems unlikely that the Mississippi legislature, or any other state legislature, would have the authority to do this either. While the Supreme Court has not yet analyzed the Mississippi RFRA, or any other state RFRA that is similar to Mississippi's, it seems logical to conclude that these state laws would not be deemed a constitutional use of state legislative power. They go beyond what the Court has previously considered appropriate and do not mirror the federal law closely enough. In addition to Mississippi, states such as Oklahoma, South Carolina, Tennessee, Louisiana, and Virginia have religious freedom laws that ignore the Court's guidance and define free exercise using the First Amendment.

Another way in which several state religious freedom laws deviate from the standard set forth by the federal RFRA has to do with the definition of the word "person" within these laws. The federal RFRA does not explicitly define person at all when it describes who the law applies to (42 U.S. Code § 2000bb-1). The Supreme Court has interpreted the federal RFRA to include protections for individuals as well as certain businesses, such as, nonprofits, religious organizations, and closely-held for-profit corporations. So far, the application of the federal RFRA has been limited to these specific entities (*Hobby Lobby*, 2014). Several state legislators, however, decided to explicitly define the term "person" in state laws, often times expanding the definition beyond that of what the Supreme Court has deemed appropriate. For example, the South Carolina Religious Freedom Act states that the term person "includes, but is not limited to, an individual, corporation, firm, partnership, association, or organization" (S.C. Code §1-32-10). The South Carolina law would therefore, in theory, apply to any type of business, whether that business is closely-held or a huge multinational corporation. The Supreme Court stated in *Hobby Lobby* that its ruling was not an implication that the federal RFRA should apply to other types of business besides closely-held corporations (*Hobby Lobby*, 2014). The South Carolina law and others like it, therefore, take a huge and somewhat risky step in including all businesses in their applications. This would essentially afford any business, even a publicly held, non-religious organization with thousands of shareholders, the same free exercise rights as any individual. The Court has never expanded the application of free exercise protections to organizations of that size and scope, and it would be dangerous if they were to do so. As Justice Ginsburg's dissent in *Hobby Lobby* speculated, it would be exceedingly difficult to determine which shareholders' religious beliefs should dictate the organization's operations. Moreover, this broad application to all businesses further opens the door to potentially discriminatory practices, especially if businesses do not wish to hire certain groups based on a religious belief.

In addition to South Carolina, states such as Indiana, Kansas, and Louisiana have expanded the definition of “person” within their religious freedom laws to include a broad array of organizations and businesses. Beyond a growing national concern that these laws may cause discrimination, there is also a concern that this broad application of free exercise protections could actually violate the Establishment Clause of the First Amendment. This part of the First Amendment states that “Congress shall make no law respecting an establishment of religion” (U.S. Const. amend. I.). This is an inherently important part of the founding fathers’ intentions when creating the U.S. Constitution and the Bill of Rights. No one religion is meant to dictate the laws of the United States or the practices of the people and businesses that operate there. In expanding the application of religious freedom laws to any kind of business or corporation, no matter the size and scope of these organizations, the state legislatures are allowing for the potential that a single religion will dictate the lives of large groups of business stakeholders, even if those stakeholders do not share the religious beliefs of the business owners. This becomes especially dangerous in the context of a business that has no official religious affiliation. Employees of a non-religious business will not likely have joined the workforce of that business with religious affiliation in mind. Therefore, a conflict arises if a business chooses to operate under the religious beliefs of a few owners, despite other owners and potentially thousands of employees sharing in the interest of the operations of that business, but not necessarily sharing in those religious beliefs. To afford large corporations religious protections equal to that of individuals would essentially force all employees, owners, managers, and stakeholders of that business to also comply with the religious preferences of the owners of that corporation. This is a dangerous path for free exercise protection to take. State religious freedom laws that expand their application so broadly could, under certain contexts, violate the intentions and principles behind the Establishment Clause of the First Amendment (*City of Boerne*, 1997; Stevens, J., concurring). While this has not yet been observed in any case law, it is certainly a concern to keep in mind as courts continue to interpret these state religious freedom acts.

Another significant change in one particular state religious freedom law is the potential for that law to be used in private lawsuits between individuals, rather than exclusively in lawsuits against the government. The Indiana Religious Freedom Restoration Act, one of the most recently passed state religious freedom laws, greatly expands the scope of how religious freedom laws can be used. In addition to expanding the definition of “person” to include organizations, and businesses, Indiana also added a provision to its law that has never before been seen in any religious freedom act (2015 SB 101). This provision states:

“A person whose exercise of religion has been substantially burdened, or is likely to be substantially burdened, by a violation of this chapter may assert the violation or impending violation as a claim or defense in a judicial or administrative proceeding, regardless of whether the

state or any other governmental entity is a party to the proceeding. If the relevant governmental entity is not a party to the proceeding, the governmental entity has an unconditional right to intervene in order to respond to the person's invocation of this chapter" (2015 SB 101).

This seems to indicate that the Indiana RFRA could be used in private, citizen suits where a party believes religious freedom is being violated, whether or not it is the government or a law that is causing the burden on free exercise. This is far beyond what the federal RFRA has ever been used for, and could once again bring free exercise protections down a dangerous path. Not only does this open the door to frivolous and numerous citizen suits, but any individual could use this provision as a defense for discriminatory behavior, whether or not there are government laws at stake. The state of Indiana received significant backlash for this and other aspects of their RFRA, with widespread outcry across the nation. Opponents of the law felt that this citizen suit provision, as well as the expanded application of the Indiana RFRA, made the law targeted toward the LGBT community (Steinmetz, 2015). In fact, this widespread outcry caused legislators in Indiana to reconsider the law, and they ultimately enacted a major change that only one other state has implemented so far: they adopted an explicit civil rights exemption within the language of the law.

## VI. State Religious Freedom Laws

### *CIVIL RIGHTS EXEMPTIONS*

The Texas and Indiana religious freedom laws are now unique in that they are the only two that have civil rights provisions intending to prevent the use of the laws for discriminatory purposes. Not even the federal RFRA has a provision that serves this purpose. The Texas civil rights provision within its RFRA is a bit broad—it merely states that “this chapter does not establish or eliminate a defense to a civil action or criminal prosecution under a federal or state civil rights law” (Tex. Civ. Prac. & Remedies Code §110.001). So ultimately, the Texas RFRA cannot be used to violate an already-existing federal or state civil rights law. However, it does not necessarily help groups that are not expressly protected by federal and state civil rights laws. There are no current federal protections in place for members of the LGBT community, and Texas does not have state protections for this group in place either (ACLU, 2015). Therefore, although many other groups will be protected from a discriminatory application of the Texas RFRA, gay, lesbian, and transgender individuals will not necessarily be afforded protection by Texas's civil rights provision.

In contrast, the Indiana civil rights amendment to the RFRA goes a step further than the Texas provision. The Indiana amendment states that the law does not “authorize a provider to refuse to offer or provide services, facilities, use of public accommodations, goods, employment, or housing to any member or members of the general

public on the basis of race, color, religion, ancestry, age, national origin, disability, sex, sexual orientation, gender identity, or United States military service” (2015 SB 50). This provision expressly prohibits any application of the Indiana RFRA that would discriminate on a wide basis of factors, including sexual orientation and gender identity. Indiana does not even have state-wide laws that prohibit discrimination on the basis of these two factors, so this RFRA amendment is a huge step for the state. This provision will likely mitigate the impact of the Indiana law’s expanded application to businesses and private citizen suits. While there are still aspects of the Indiana RFRA that go far beyond what the federal law and Supreme Court have protected in the past, the explicit civil rights provision is an example of steps that states can take to balance religious freedom with equal treatment.

### DISCRIMINATION

While Texas and Indiana have taken steps to attempt to balance both religious freedom and anti-discrimination protections, other states have taken the opposite approach, instead choosing to adopt provisions to their religious freedom laws that would explicitly target certain groups or practices. Extreme measures within certain state religious protection laws tend to favor certain religious beliefs, while targeting others. This tendency defeats the entire intended purpose of the original RFRA, which is to provide equal protection to all religions.

The Florida Religious Freedom Restoration Act is an example of a state law that has added a targeted provision to its language. The Florida law has a provision that expressly allows any religious organization or individual authorized to solemnize a marriage to refuse to do so if the marriage in question is not representative of that individual or organization’s religious beliefs (Fla. Stat. §761.01). It is fairly obvious that this provision of the Florida law expressly grants permission to any individual authorized to certify a marriage to refuse marriage rights to gay or lesbian couples. This clearly supports the beliefs of religious sects that disagree with the concept of same-sex marriage, while targeting those individuals whose do not share in those beliefs. Expressly allowing individuals to deny marriage rights to LGBT individuals could be considered discriminatory, and may create an interpretation that the Florida RFRA is targeting the rights of LGBT individuals. The fact that Florida’s law has this provision removes any degree of neutrality from the Act. Florida’s RFRA does not equally protect all religious beliefs, but rather caters to the beliefs of a particular group of religions with a certain viewpoint. It is also unclear as to whether this marriage provision in Florida’s law could be used to offset the Supreme Court’s ruling in *Obergefell v. Hodges*, which legalized same-sex marriage under the Fourteenth Amendment (*Obergefell*, 2015). The federal RFRA was created with the purpose of offsetting the unpopular Supreme Court ruling in *Employment Division v. Smith*, so it is possible that state courts may interpret a state RFRA as being able to offset the *Obergefell* ruling. Until more case law exists under the Florida RFRA in the aftermath of *Obergefell*, it is unclear how the applica-

tion of the law will play out. It is clear, however, that the Florida RFRA goes farther than the federal RFRA does, and creates the potential for discriminatory applications.

Even though not all religious freedom laws have measures as extreme as those cited in the Florida law, in theory any religious freedom law that does not have an explicit civil rights exemption can be used in a discriminatory application. While there are no federal laws that protect against discrimination on the basis of sexual orientation and gender identity, twenty-three states have passed laws that do protect against discrimination based on these factors (ACLU, 2015). Additionally, there are some city-specific and local ordinances that protect the LGBT community from discrimination. It is entirely possible that any religious freedom act without a civil rights amendment could be used to invalidate any of these anti-discrimination laws on the basis of free exercise protection. While it has been argued by supporters of religious freedom laws that this concern is unwarranted, as civil rights laws would pass the compelling interest test, this has not in fact been the case.

Courts in both Michigan and Kentucky have used both the federal and a state religious freedom law to examine anti-discrimination laws under strict scrutiny, and in both instances the laws failed the test. There is another case pending in Georgia, in which an anti-discrimination ordinance will be examined under the Religious Freedom Restoration Act (Ford, 2015). If civil rights laws cannot pass this super strict standard of review, then it is clear that religious freedom laws, at both the state and federal level, can create a mechanism for discrimination against these groups. The Supreme Court has yet to rule on an application of a religious freedom law that nullifies the validity of an anti-discrimination law. However, until it does, it appears that the nation-wide concerns about the potentially discriminatory applications of state and federal religious freedom restoration laws are warranted. This troubling conflict between civil rights and religious freedoms in the United States has no clear resolution, and courts and legislators will surely struggle to balance these two interests going forward.

## **VII. Implications for the Future and Recommendations**

The outcome of the continuing conflict between religious freedom laws and civil rights protections in the United States will largely depend on future rulings on the issue by the Supreme Court. The Court has clearly modified its approach to free exercise protections over the years, and it will very likely continue to do so. Until the Court decides to hear a religious freedom case that directly deals with a discrimination law, states will continue to rule on these laws in ways that are consistent with their distinct religious freedom acts. However, there are measures that, if implemented, could help balance the conflict between free exercise and equal protection.

While the civil rights provision of the Indiana Religious Freedom Restoration Act was implemented due to public outrage regarding the extremeness of the original version of the law, the final version shed light onto new possibilities. By expressly forbidding the use of the RFRA for discriminatory purposes, and stating that the law cannot

be used to deny services to any individual on the basis of sexual orientation and gender identity, the law has effectively mitigated any effect it would have had on those minority groups. Since the implementation of the civil rights provision, there have been no discriminatory RFRA cases resulting from the Indiana law. Meanwhile, in Kentucky, where the state RFRA does not have a civil rights exemption, the RFRA was used to strike down the application of the Lexington Fairness Ordinance to the for-profit company Hands on Originals (*Hands on Originals*, 2015). This application of the Kentucky RFRA could have been prevented if there had been a civil rights provision in the state law as well.

The federal RFRA, as well as all twenty-one state religious freedom laws, could benefit from the addition of an explicit civil rights amendment, much like the one that was implemented by the state of Indiana. An amendment such as this would allow legislators the additional protection for free exercise that they have aimed for through the implementation of religious freedom acts, while also ensuring that the laws are not used for discrimination. Civil rights provisions would also somewhat mitigate the expanded application of religious freedom acts to businesses. Because there are currently no nation-wide protections for members of the LGBT community, a civil rights provision that explicitly prevents the use of a religious freedom law to deny services to this group would be a huge step toward ensuring equal protection in the United States. While the civil rights provision in the Texas RFRA was a step in the right direction, this particular protective measure is not as effective as the amendment implemented in Indiana. Because the Texas provision only prevents the RFRA from violating any already-existing civil rights laws, it does not expressly protect groups like the LGBT community. Many states do not have civil rights laws that protect gay, lesbian, and transgender individuals from discrimination, and no federal laws do this either. Therefore, the Texas RFRA could still allow for discrimination.

The Indiana law, on the other hand, expressly lists the groups for which it prohibits discrimination, thereby making it irrelevant as to whether the state already has certain civil rights laws in place. The best step that Congress and the states could take to prevent discrimination and mitigate some of the lack of neutrality that exists in certain religious freedom laws would be to implement civil rights amendments similar to that of Indiana. In doing so, religious freedom laws could be used for non-discriminatory reasons, in cases where a neutral and generally applicable law is truly creating a substantial burden on a religious belief. However, with a civil rights exemption, these laws could no longer be used as a mechanism for discrimination under the guise of religious freedom.

In the end, there is no perfect solution to balancing free exercise protection with civil rights protection. The Supreme Court has struggled with determining how much protection the First Amendment grants to free exercise, and whether or not legislators have the right to increase the protection offered by the Free Exercise Clause. The Court has also struggled to determine whether free exercise protections apply to individuals

alone, or to organizations and businesses as well. The Court's decisions regarding free exercise have evolved significantly over the years, and they will need to evolve further as the issue of balancing this protection with civil rights for minority groups continues to attract national attention. If the Court were to expressly rule that the RFRA, and other laws like it, could not be used to discriminate, then states would need to adjust their laws accordingly. However, until the Court rules on this issue, state courts can interpret free exercise protection and religious freedom acts as they wish. Hopefully, states will start to follow Indiana's example, and adjust their laws in a way that is consistent with equal protection for all.

## References

- ACLU.** Non-Discrimination laws: state by state information - map. Retrieved October 25, 2016, from <https://www.aclu.org/map/non-discrimination-laws-state-state-information-map>
- Alabama Religious Freedom Amendment,** Ala. Const. Art. I, §3.01 (1998).
- Alstyne, W. W. (1996).** "The failure of the Religious Freedom Restoration Act under Section 5 of the Fourteenth Amendment". *Duke Law Journal*, 46(2), 291-306. doi:10.2307/1372958
- Ariz. Rev. Stat.** §41-1493.01 (1999).
- Blatnik, E. J. (1998).** "No RFRAF Allowed: The Status of the Religious Freedom Restoration Act's Federal Application in the Wake of *City of Boerne v. Flores*." *Columbia Law Review*, 98(6), 1410-1425. doi:10.2307/1123302.
- Burwell v. Hobby Lobby,** 134 S. Ct. 2751 (2014).
- Burwell v. Hobby Lobby,** 134 S. Ct. 2751 (2014) (Ginsburg, J., dissenting).
- Cantwell v. Connecticut,** 310 U.S. 296 (1940).
- Church of the Lukumi Babalu Aye v. City of Hialeah,** 508 U.S. 520 (1993).
- City of Boerne v. Flores,** 521 U.S. 507 (1997).
- City of Boerne v. Flores,** 521 U.S. 507 (1997) (Stevens, J., concurring).
- Conn. Gen. Stat.** §52-571b (1993).
- Employment Division, Department of Human Resources of Oregon et al. v. Smith et al.,** 494 U.S. 872 (1990).
- Equal Employment Opportunity Commission v. R.G. & G.R. Harris Funeral Homes, Sean F. Cox, Inc.,** No. 14-13710 (E.D. Mich. 2016).
- Fla. Stat. §761.01,** et seq. (1998).
- Ford, Z. (2015).** "Court Says 'Religious Freedom' Gives T-Shirt Company The Right To Discriminate Against LGBT Group." Retrieved October 15, 2016, from <https://thinkprogress.org/court-says-religious-freedom-gives-t-shirt-company-the-right-to-discriminate-against-lgbt-group-cac86c5ba57#.mqy82xopx>

**Gonzales v. O Centro Espirita Beneficente Uniao do Vegetal**, 546 U.S. 418 (2006).

**Hamilton, M. A. (2015)**. “The Case for Evidence-based Free Exercise Accommodation: Why the Religious Freedom Restoration Act is Bad Public Policy.” *Harvard Law & Policy Review*, 9(1), 129-151.

**Hands on Originals, Inc. v. Lexington-Fayette Urban County Human Rights Commission**, No. 14-CI-04474 (Fayette Cir. Ct. Apr. 27, 2015).

**Harvard Law Review Association. (2012)**. “Equal Protection — Sexual Orientation — First Circuit Invalidates Statute That Defines Marriage as Legal Union Between One Man and One Woman.” *Harvard Law Review*, 126(2), 611-618.

**Indiana Religious Freedom Restoration Act Amendments, 2015 SB 50 (2015)**.

**Kan. Stat. §60-5301**, et seq. (2013).

**Ky. Rev. Stat. §446.350** (2013).

**Idaho Code §73-402** (2000).

**Lane-Steele, L. (2015)**. “Masking Discrimination: How the ‘Second Wave’ of RFRAs Can Weaken Protections for LGB Individuals.” *National Lawyers Guild Review*, 72(2), 109-121.

**Lipka, M., & Wormald, B. (2016, February 29)**. “How Religious is Your State?” Retrieved December 19, 2016 from <http://www.pewresearch.org/fact-tank/2016/02/29/how-religious-is-your-state/?state=alabama>

**Luchenitser, A. J. (2015)**. “A New Era of Inequality? Hobby Lobby and Religious Exemptions from Anti-Discrimination Laws.” *Harvard Law & Policy Review*, 9(1), 63-88.

**Lyng v. Northwest Indian Cemetery Protective Ass’n**, 485 U.S. 439 (1988).

**Maatman, G. L., & Karasik, A. W. (2016, August 19)**. “EEOC loses landmark transgender discrimination case.” Retrieved November 30, 2016, from <http://www.workplaceclassaction.com/2016/08/eoc-loses-landmark-transgender-discrimination-case/>

**Mississippi Religious Freedom Restoration Act**, Miss. Code §11-61-1 (2014).

**Montanaro, D. (2015, April 1)**. “Indiana Law: Sorting Fact from Fiction from Politics.” Retrieved September 1, 2016, from <http://www.npr.org/sections/itsallpolitics/2015/04/01/395613897/sorting-fact-from-fiction-from-politics-on-the-indiana-law>.

**New Mexico Religious Freedom Restoration Act**, N.M. Stat. §28-22-1, et seq. (2000).

**NCSL. (2015, September 03).** 2015 state religious freedom restoration legislation. Retrieved September 8, 2016, from <http://www.ncsl.org/research/civil-and-criminal-justice/2015-state-rfra-legislation.aspx>

***Obergefell v. Hodges***, 135 S. Ct. 2584 (2015).

**Oklahoma Religious Freedom Act**, Okla. Stat. tit. 51, §251, et seq. (2000).

**Preservation of Religious Freedom Act**, La. Rev. Stat. §13:5231, et seq. (2010).

**Protection of Religious Exercise in Land Use and by Institutionalized Persons**, 42 USCS § 2000cc-5 (2000)

**Religious Freedom Protection Act, Pa.** Stat. tit. 71, §2403 (2002).

**Religious Freedom Restoration Act, Ill.** Rev. Stat. Ch. 775, §35/1, et seq. (1998).

**Religious Freedom Restoration Act, Mo.** Rev. Stat. §1.302 (2003).

**Religious Freedom Restoration Act, R.I.** Gen. Laws §42-80.1-1, et seq. (1993).

**Religious Freedom Restoration Act, 2015 SB 975** (2015).

**Religious Freedom Restoration Act, 42 U.S.C.** §§ 2000bb-2000bb-4 (1993).

**Religious Freedom Restoration Act of 1993** (1993 - H.R. 1308). Accessed April 17, 2016. <https://www.govtrack.us/congress/bills/103/hr1308>.

**Religious Freedom Restoration, 2015 SB 101** (2015).

**Ryan, J. E. (1992).** “Smith and the Religious Freedom Restoration Act: an Iconoclastic Assessment.” *Virginia Law Review*, 78(6), 1407-1414. doi:10.2307/1073457.

**Seeger, S. C. (1997).** “Restoring Rights to Rites: The Religious Motivation Test and the Religious Freedom Restoration Act.” *Michigan Law Review*, 95(5), 1472-1486. doi:10.2307/1290013

***Sherbert v. Verner et al., Members of South Carolina Employment Security Commission, et al.***, 374 U.S. 398 (1963).

**Siegel, S. A. (2006).** “The Origin of the Compelling State Interest Test and Strict Scrutiny.” *The American Journal of Legal History*, 48(4), 355-407. doi:10.2307/2546981.

**South Carolina Religious Freedom Act**, S.C. Code §1-32-10, et seq. (1999)

**State Religious Freedom Restoration Acts.** (2015, October 15). Retrieved November 12, 2016, from <http://www.ncsl.org/research/civil-and-criminal-justice/state-rfra-statutes.aspx>.

***State of Wisconsin v. Jonas Yoder, Wallace Miller, and Adin Yutzy***, 406 U.S. 205 (1972).

**Steinmetz, K. (2015, March 30).** “What is Indiana’s ‘Religious Freedom’ Act Really About?” Retrieved September 2, 2016, from <http://time.com/3764347/indiana-religious-freedom-discrimination-act/>

**Tenn. Code §4-1-407** (2009)

**Tex. Civ. Prac. & Remedies Code §110.001**, et seq. (1999)

**U.S. Const. Amend. I.**

**Utah Code Ann. § 63L-5-101 to -403** (2008)

**Va. Code §57-2.02** (2007)

**Wormald, B. (2015).** “Religious Landscape Study.” Retrieved November 12, 2016, from <http://www.pewforum.org/religious-landscape-study/region/south/>

**42 U.S. Code Chapter 21B** – Religious Freedom Restoration. LII / Legal Information Institute. Accessed April 17, 2016. <https://www.law.cornell.edu/us-code/text/42/chapter-21B>.

## MANAGING STIGMA: WOMEN DRUG USERS AND RECOVERY SERVICES

By Nayeong Lee and Miriam Boeri\*

*Women who use drugs are stigmatized for their drug use behavior, which marginalizes them from mainstream society. Stigmatization can be viewed as an attempt by social services to exert control. Research shows that these strategies do not work well for discouraging drug use; whereas attempts to reduce the stigma related to drug use can encourage users to stop use. Using qualitative methods and grounded theory analysis, the goal of this study is to examine (1) the stigmatization of drug use through different stages; (2) how stigmatized women drug users perceive normality; and (3) barriers and challenges to recovery. Based on in-depth interviews from 20 women who used methamphetamine, the analysis focuses on stigmatization before the initiation of drug use, difficulties related to stigma as drug users, and challenges due to stigmatization as they recover from drug use. Findings show that women are stigmatized before they use drugs, face more stigma as they use, and even during recovery society still holds onto the label of former drug user, making it difficult to avoid stigma. The findings contribute to a better understanding of how stigmatization of women drug users impacts their recovery and provides suggestions for social service and treatment providers.*

Keywords: drug use; drug treatment services; stigma; normality; social control.

### I. Introduction

Although many women initiate the use of drugs as a way to self-medicate and address social pressures, they are stigmatized by society for using drugs, and women who are mothers face even greater stigmatization (Lende, Leonard, Sterk, & Elifson, 2007; Radcliffe, 2011; Roberts, 1991). This study explores the initiation of drug use among women, as well as the difficulties women face as drug users and as they try to recover from drug use. The goal of this study is to examine (1) the stigmatization of drug use through different stages; (2) how stigmatized women drug users perceive normality; and (3) barriers and challenges to recovery. This qualitative analysis of the life histories of 20 women who used illegal drugs focuses on their social roles, the influences of initiating and using drugs, and the process of recovering from a stigmatized social identity. This study is significant in that stigmatization is examined at each stage of

---

\* The project described was supported by the National Institute on Drug Abuse, Award numbers 1R15DA021164; 2R15DA021164. The content is solely the responsibility of the authors and does not necessarily represent the of official views of the National Institute on Drug Abuse or the National Institutes of Health. Miriam Boeri, PI mboeri@bentley.edu.

drug use with an in-depth focus on gendered social roles. The findings contribute to a better understanding of how stigmatization of women drug users impacts their recovery. The conclusions provide suggestions for social service and treatment providers.

The analytical framework for the study is informed by the concepts of stigma and normality (Goffman, 1963; Neale, Nettering, & Pickering, 2011; Radcliffe, 2011). The women who shared their stories were mostly poor or working-class white women living precariously on the edge of suburban middle-class society. The analysis distinguishes stages of drug use revealed in the women's narratives. During the period called the *before* stage, various forms of stigmatization made the women feel insecure, which is one of the reasons to why they turn to drugs. However, to conceal their use and avoid stigma, the women try to become what they perceive as "normal" to meet gendered expectations set by society. The women face more stigmatization once they start using drugs (the *during* stage) due to society's perceptions of drug users. The *after* stage consists of the period when women stop using drugs and their efforts towards becoming "normal" while the recovering from drug use. Findings show that although some women drug users make the effort to change, society still holds onto the label of "former drug user," making it difficult to avoid stigma.

How do women in a stigmatized status maintain self-esteem? How do they fit in with mainstream society while using illegal drugs? After they are exposed as being drug users, how do they live with the identity of a former drug user? What types of treatment and resources are available for the women drug users to recover? Are these services effective? How does being a drug user or recovering drug user affect their mother roles for those who have children? How do perceptions of normality impact each stage of their drug use?

To answer these questions, the analysis focuses on the women's perceptions of stigma and normality as they tried to "fit in" society and perform the gendered social roles at the standards they believed were expected of them. This study also examines the different types of social control that influenced the women, including treatment services, religion, and relationships, as well as the extent that gender played in women's social lives.

## II. Background

The conceptual framework of stigma and normality frames the analysis of the women's lives. The literature on the impact of gender on drug use and access to services informs the analysis of their recovery efforts.

### *STIGMATIZATION*

Society has certain perceptions of what is normal and what is not. People who are considered different from "normal" are stigmatized, perceived as deviant, and often marginalized. White (2002) explains stigma as the process of labeling, stereotyping, social rejection, exclusion and extrusion, as well as the internalization of community

attitudes in the form of shame by person and family. The literature on stigma draws primarily from the work of Erving Goffman (1959, 1963).

Goffman distinguished between those who are “discredited” when their stigma is known, from those who are “discreditable” when they are able to conceal their perceived stigma. Both discredited and discreditable stigma are negotiated through “impression management” (Goffman, 1959). Individuals with hidden stigma “may expend much energy to ensure that stigma-related ‘leakages’ do not occur” (Pachankis, 2007, 335). Yet, while they try to fit into “normal” or conventional society, they also struggle with feelings of insecurity, isolation and anxiety (Hetrick & Martin, 1987).

According to Goffman (1963), stigmatized individuals try to prove themselves by drawing attention to more positive aspects of their identity while attempting to conceal the aspect that is stigmatizing. Stigmatization can make individuals feel insecure, and they often turn to those similar to them (other stigmatized individuals) for social support (McKenna & Bargh, 1998). Goffman (1963) further explains the limitations and boundaries of what actually defines stigmatization as well as the underlying social functions of stigmatization, such as social control.

Studies consistently show that stigma negatively impacts health and contributes to health disparities found among marginalized populations (Chaudoir, Earnshaw, & An-del, 2013). Moreover, research shows that even hidden stigma can negatively impact psychological well-being and lead to social isolation and alienation (Pachankis, 2007; Quinn & Chaudoir, 2009).

Similarly, White (2002) discusses how social stigma toward alcohol and other drug addiction may be an obstacle to resolve problems or to even come up with a strategy to solve the issue of addiction. White (2002) further identifies several myths surrounding stigma and how it is important to understand the difference between stereotypes and facts. Moreover, stigma used by social services as a control mechanism limits attempts at recovery.

#### *SOCIAL CONTROL AND SOCIAL SERVICES*

Social control theory proposes that delinquency and criminal behavior are a consequence of changes in the quality and strength of social ties (Laub & Sampson, 2003). Individuals are more likely to engage in deviant behaviors, such as drug use, when their bonds to society are weak or broken. Formal social control of behavior, such as laws and bureaucratic rules, are distinguished from informal social controls. Informal social control acts through the bonding that comes with attachment to others and ties to mainstream social institutions, such as school, work, and religious affiliation. Informal social control works through “strong bonds with family, friends, work, religion, and other aspects of traditional society motivating individuals to engage in responsible behavior or [acceptable] social norms” (Moos, 2007).

Being stigmatized marginalizes the discredited individual from mainstream society, and stigmatization by social services is often viewed as an attempt to exert social control (Palamar, Halkitis, & Kiang, 2013). Focusing on the stigmatization of drug us-

ers within society, Radcliffe and Stevens (2008) examine the categorization of “junkies,” which refers to long-term drug users. The authors find that “junkies” are stigmatized even within treatment services. While studies consistently show that treatment clients receive real or perceived stigma from the staff and professionals who work with them (Boyd, 1999; Campbell, 2000; Stengel, 2014), research in other healthcare contexts shows that stigma can be mitigated by acts of kindness (Walter et al., 2015).

Studies find that labeling, stereotyping, and discrimination are commonly used to stigmatize (Link & Phelan, 2001; Kyons et al., 2015; White, 2002), yet research shows that these strategies do not work well for discouraging drug use and are associated with adverse mental and social health outcomes (Bayer, 2008; Palamar, 2012; Radcliffe & Stevens, 2008). Jiménez et al. (2011) argue that stigmatizing actions are thought to uphold social order, but in fact they cause social divisions. Since drug users can lose family, friends, employment, housing, school loans and other social and economic benefits, many attempt to conceal their use even from those who can help them (Palamar, 2012). In contrast, attempts to reduce the stigma related to drug use and eliminate discriminatory attitudes toward drug users, particularly by health professionals and law enforcement, can encourage users to seek needed care (Rivera et al. 2015).

### Normality

In order to understand stigmatization, it is important to define what is considered to be “normal.” Society has set perceptions of how individuals should act and standards of normality. However, it is difficult to define normality as there are varying opinions of what “being normal” actually means. For the purposes of this study, drug users are seen as deviant from mainstream society standards for normal, but even within the drug user community, definitions of normal vary.

Copes, Hochstetler and Williams (2008) explore the concept of identities within the drug user communities. According to these authors, social identity is referred to as a multivalent process where individuals identify themselves in terms of being similar to some people and different to others. Individuals construct boundaries and identities that separate them from others who they view as having lower status. Although society assigns negative labels for drug users, their identity is also influenced by other drug using individuals who define distinctions between drug users. For example, Copes and his colleagues explain that “hustlers” (i.e., minor criminals) within the drug community believe that they are above “junkies” (i.e. hardcore criminals). Although “hustlers” are still part of the same drug community as “junkies,” they see themselves as a step closer to being “normal” than “junkies,” who have fallen the lowest levels of drug user category.

Other research shows that even in the drug-using community, an occasional or weekend user is viewed in less negative terms than an addict or “junkie,” since they appear to control their use (Boeri 2004, Radcliffe & Stevens, 2008). People in drug treatment may attempt to navigate stigma renegotiating what normal means to them (Nettleton, Neale, & Pickering, 2012). Others try to reduce multiple stigmas by avoiding treatment itself, which is associated with the most stigmatized drug using stereotypes (Radcliffe & Stevens, 2008).

*PERCEPTIONS OF WOMEN DRUG USERS*

Gendered perceptions of drug users exist in both mainstream and drug-using worlds. Goffman (1963) explains how the stigma of group identity is related to the stigma of race, nation, and religion, affecting a whole group rather than an individual. Women as a gendered group face greater stigmatization than men for using drugs since they go against the character traits of perceived female identity. The stigma of drug use is also greater for mothers since they are expected to be the caregivers, raise children, and be more family oriented than fathers. Women are usually aware of these double standards and try to present a good image to society by hiding their drug use if they are mothers (Goffman, 1959).

Gendered differences in perceptions of substance use can be seen in the use of both legal and illegal drugs. Females are more negatively viewed when they smoke in public compared to males. While females who smoke are seen as “trash” and “sluts,” males who smoke are seen as “more masculine” and “attractive” (Nichter, 2006, 112). Therefore, females tend to smoke in groups in a more hidden environment to preserve their good reputation while males are able to freely smoke in public. This can be translated into the perception of how women are viewed when they use stronger drugs, such as cocaine or methamphetamine.

Substantial research shows that female drug users face stigmatization as “bad women” because they violate gender-role expectations (Boeri, 2013; Campbell, 2000; Ettore, 1992). Women who use drugs while pregnant are identified as having a moral ‘failings’ and ascribed a “spoiled identity” (Stengel, 2014; Stone, 2015). Beyond the stigma of being a drug-user, women can feel the burden of multiple stigmas, such as poverty, minority status, unemployment, transgender identification, and older age (Connera & Rosen, 2008; Lyons et al., 2015; Roberts, 1991).

Dluzen and Liu (2008), exploring the differences between male and female users of methamphetamine, found that while violence and excessive use were more prevalent behaviors among female than male methamphetamine users, women responded to treatment better than males. Yet, women users have limited access to treatment resources and face more challenges during recovery than do men (Boeri, Gardner, Gerken, Ross & Wheeler, 2016; Maher & Hudson, 2007).

*BARRIERS TO RECOVERY*

The greater stigmatization of female drug users presents several barriers for women when they try to get their lives back to “normal,” often without sufficient support, resources, or guidance. Luck, Elifson, and Sterk (2004) found that women who are drug users and are on the welfare system are perceived negatively by others. There are many people who believe that women who use drugs and receive the welfare system are “undeserving poor,” and some people in society may believe that the welfare system is funding the dependence on drugs. There are also speculations that the use of

drugs leads individuals to welfare dependence, but the authors show that this is a more complex relationship than assumed. Although women in the welfare system are willing to take responsibility of their lives, the negative views from society make them feel disrespected and reinforces feelings of powerlessness.

The services needed are not as accessible as many think due to the stigmatization of those who apply for aid. Boeri, Tyndall, and Woodall (2011) show how the barriers to services that women need to help themselves improve their lives can be a reason why women use or continue to use drugs. Their study found that many of the women who are addicted to methamphetamine are willing to seek help, but barriers preventing the women from moving forward create a cycle of defeat. For example, the lack of identification (e.g., ID card) created barriers getting into shelters, being homeless created a barrier to obtaining an ID card, and without identification, these women were unable to apply for the programs that could help them. When social services turned them down because they did not meet all the eligibility criteria, the women found it more difficult to meet social expectations without access to the needed resources (Boeri, 2013).

Focusing on resources, Hall, Baldwin, and Prendergast (2001) propose that more treatment solutions are needed. They explain how community-based residential treatment programs are key elements in increasing the likelihood to avoid relapse. However, women often faced difficulties accessing these programs, ultimately leading them back to drugs or jail. These barriers exist on all levels. Many women are very limited in resources even though society claims that there are a multitude of services available to them. Some of these women have no means of obtaining the resources in their current social situations. For example, women who live in areas with no public transportation cannot get to needed services, and women who do not have a phone cannot make appointments or receive a call back when they leave messages (Boeri, 2013; Luck et al., 2004). Therefore, even though programs exist, it is also important to consider access to them for disadvantaged women (Sered, 2014; Woodall & Boeri, 2013). The negative perception of recovering or “former drug users” is another reason that these women do not go to these programs.

The literature on the social impact of stigmatization and perceptions of being normal discussed above suggests that there is greater stigmatization of female drug users compared to men who use drugs. It also shows that negative perceptions of drug use exist in mainstream as well as drug-using worlds, and these perceptions often create barriers for women attempting to recover. Informed by the literature, this study examined the lives of 20 women who used methamphetamine and other illegal drugs with the goal to understand every stage of their drug use, and with particular focus on barriers to recovery. Using a qualitative method known as “grounded theory” (Charmaz, 2006), the study examined how female drug users begin using drugs and how they strive to achieve a sense of normality while facing stigmatization as current or former drug users.

### III. Methods

The data used for the analysis in this paper was drawn from a larger study on women who use methamphetamine conducted by Boeri (2013).<sup>1</sup> Female methamphetamine users were drawn from the suburban counties outside a large city in southeastern USA. Participants were recruited using a combination of snowball, targeted and theoretical sampling methods (Glaser & Strauss, 1967; Strauss & Corbin 1998; Watters & Biernacki 1989). Snowball sampling, also called chain referral, involved asking participants and interested inquirers to refer another potential participant to the study. Targeted sampling involved ethnographic fieldwork in communities where drug use was prevalent and establishing relationships with community members to reach potential participants. Fliers were used to publicize a “methamphetamine study” with a study number for interested individuals to call for more information. Once potential participants for the study were located or made contact through the study phone number, Boeri discussed the study time commitment, how the interview would be conducted, anonymity and confidentiality issues, and reimbursement for their time. Theoretical sampling, used in grounded theory, involved the collection of data based on theory that emerged from the data while it was collected, designed “to maximize opportunities to discover variations among concepts and to densify categories in terms of their properties and dimensions” (Strauss & Corbin, 1998, p. 201). Theoretical sampling was used to insure a diversity of participant experiences, as well as to focus on recruitment of specific kinds of users that emerged from ongoing analysis of the data.

While the study design seems straightforward, the reality of conducting research on illegal and stigmatized behaviors, especially in the suburbs, has aspects that can only be learned by doing it. People who used methamphetamine do not usually “hang-out” during work hours. Ethnographic fieldwork conducted by Boeri and her research team typically involved going out all day to find field sites, distribute fliers, and talk to anyone interested. In the evening and night, they frequented bars, clubs, and all night diners.<sup>2</sup>

A screening process helped to ensure that interested participants passed the eligibility criteria to participate in the study. Eligibility criteria included having used methamphetamine for at least six consecutive months, living in the suburbs, and being at least 18 years old. After a participant was screened and consented, the interview was conducted in a safe location agreed upon by the interviewer and participant. Places used included the interviewer’s car, the participant’s home, motel rooms, private university or library rooms, and cafes or bars during quiet hours. No identifying material, such as name and address, was collected. Data materials were identified with a study number. The interview consisted of three inter-related components: (a) a life history matrix; (b) a drug history matrix, and (c) an audio-recorded in-depth life history interview. The major themes in the interviews focused the context of drug use, time of use, interaction with others, social roles and health issues.

---

1. Boeri conducted the study on female methamphetamine users from 2008 to 2011.

2. Boeri conducted ethnographic fieldwork with a trained research team in a southern state.

The study was approved by the university's Institutional Review Board and received a "certificate of confidentiality" from a federal agency to protect the study data. The audio-recorded interviews were semi-structured and open-ended so that participants could also lead the conversation into new areas of interest. The recordings were then transcribed word-for-word.

For comparison purposes both women who had used in the last month (current users) and woman who had not used in the last month (former users) were included in the sample. Current users were defined as having used methamphetamine at least one time in the past month. Former users were defined as having used the drug for at least six consecutive months in the past but having been drug-free for the last month. Most women were polydrug users. None of the women were currently in drug treatment.<sup>3</sup>

### *SAMPLE*

The 20 participants used in the analysis for this paper were selected because they were mothers who had children in their care when they were using methamphetamine or they were of childbearing age. The mothers who currently had children in their care (custodial mothers) provided a comparison group for those who were not mothers or who no longer lived with their children. Most of the women in this sample were poor or lived in low-income communities. The life histories of these women provided rich data on life experiences and situations that could answer the key questions related to stigma, normality and challenges to recovery among female drug users.

As shown in the Table 1, ages ranged from 18 to 49, and all but one are white, which is representative of the suburban towns where the study was conducted. Six women were currently using methamphetamine at the time of the interview. All the women were using another legal or illegal substance, including eight who used a legal substance (alcohol (A) or tobacco (T)) and the remaining who were using both legal and illegal substances during the month before the interview. Among those who continued to use an illegal substance, the majority used marijuana (M) or prescription pills (P) obtained illicitly (cocaine = C). Significant to the analysis is that most of the women experienced violent abuse at some time, primarily rape and domestic violence. Seven of the women were custodial mothers at the time of the interview; among these, two were current users of methamphetamine and six had experienced violent abuse in their lifetime. All names used are pseudonyms.

---

3. See Boeri (2013) for a description of the full sample.

TABLE 1: SAMPLE DEMOGRAPHICS

Pseudonym	Age	Race/ Ethnicity	Meth Use	Other Drugs*	Abuse	Custodial Mother
Abigail	20	White	Current	T, A, C, O	N	N
Audrey	18	White	Former	T, A, C	Y	N
Beth	23	White	Former	T	Y	Y
Chloe	18	White	Former	A	Y	N
Dolly	37	White	Former	T, A	Y	Y
Emma	26	White	Current	T, A, M,	Y	N
Grace	22	White	Former	T, A, M, P	Y	N
Harper	22	White	Former	A	N	N
Isabella	27	White	Former	T	Y	N
Katy	22	White	Current	T, A, P	N	N
Lily	21	White	Current	T, A, M, P	N	N
Linda	49	White	Former	T	N	Y
Lisa	30	White	Former	M	N	Y
Lydia	45	White	Current	T, A, M	Y	N
Madeline	22	White	Former	P	Y	N
Mercedes	34	Latina	Former	T	Y	Y
Rachel	19	White	Former	T	Y	Y
Sophia	23	White	Former	T, A, M	Y	Y
Tammy	35	White	Current	T, A, P, C	Y	Y
Tiffany	25	White	Former	A	Y	N

#### ANALYSIS

The analysis involved coding over 400 pages of interview transcripts, which resulted in 70 pages of coding.<sup>4</sup> The interviews were coded using modified grounded theory methods (Charmaz, 2006, Strauss and Corbin, 1998). Whereas grounded theory derives categories or themes that emerge from the data, a modified version makes use of different concepts derived from the literature as well as the data while focusing on relations between the categories (Malterud, 2001). The coding was conducted by reading each transcript and looking for insights related to the key objectives of this paper: (1) how female drug users manage stigmatization of their drug use; (2) how female drug users reclaim a sense of normality; and (3) barriers and challenges to controlling drug use and/or recovery. These were grouped by concepts found in the literature related to the stigmatization of women who used drugs, perceptions of normality, and barriers to services (Neale et al. 2001; Boeri et al. 2011). The coding involved a two-step process. As the analysis continued, the conceptual categories were expanded to

4. The first author conducted the coding for this paper; the second author helped develop the conceptual themes and sub-themes.

include patterns found in the women's perceptions of social life, drug use and recovery efforts (Radcliffe, 2011). For example, some women perceived the benefits of some drugs helped them to stop other drug use. This was coded as a "self-help" category and added to the concept of "services" to indicate "services/help."

During the first step of coding five key concepts were developed: (1) mother/mom; (2) normal; (3) child/children/kid(s); (4) difficulty/hardship; and (5) services/help. The concept "mother/mom" was coded in the sense of women's roles as a mother rather than discussing their own mothers. This allowed greater insight on the different aspects of being a mother as well as a drug user and how the two conflicting roles were maintained. The concept "normal" was used to understand what being normal meant to the participants. This showed how the women viewed society's concept of "normal" and how they thought they measured up to that definition/standard. The categorical concept of "child/children/kid(s)" was used to determine if and when the women fulfilled their roles as mothers and how this impacted their drug use and recovery. This coding also showed the different perspectives of being a good mother as well as various methods the women used to provide for their family. The concept of "difficulty/hardship" related to the measures that the women had to take in order to survive or recover. There were difficulties that these women faced for being a drug user and a mother, so this term facilitated an examination of how they overcame the difficulties or not, and how it affected their lives. Lastly, the concept "services/help" encompassed the services provided by society as well as the women's perceptions of self-help or receiving help from other than public services. Society's help refers to the services provided by the government such as treatment programs; whereas the perception of self-help refers to perceived benefits that the women gained from doing drugs, such as weight loss, energy, and self-medication, as well as help they received from neighbors or strangers.

The second step in the coding process involved examining the coding under these main concepts to find relations between them, and developing conceptual themes for the analysis. Finding that some women used drugs to manage the stigma they felt as being poor or to cope with feelings of hopelessness before their drug use started, led to conceptualizing the stages of drug use. The theme of *Stages* includes before, during, and after drug use. The *before* stage covers the initiation of drug use and reveals why the women became dependent on the use of drugs. The *during* stage explores how drugs helped these women function in life as well as cope with stigma, enabling them to feel "normal" while using drugs. The *after* stage examines their lives after being exposed or discredited for using drugs, how the women handle the criticism and their path to normality without using drugs. These are further organized by sub-themes that were found to impact the women in each stage.

The second theme that emerged from the coding is *Social Control*, which is divided into sub-themes of *treatment services*, *religion*, and *social environment*. These sub-themes represent social control agents in the sense that they influence the women to control their drug use (in this case methamphetamine) or become drug-free. The social environment refers to the living situations of the women while trying to recover from

drug use, which included jail and drug courts (formal social control), and their own neighborhoods and social networks (informal social control). The findings describe these themes in more detail, with examples from the transcripts. The women's words are shown verbatim without editing for grammar.

## IV. Findings

### *STAGES OF DRUG USE: BEFORE, DURING, AND AFTER*

The use of drugs can be divided into several stages which do not necessarily correlate to the amount of drugs used throughout the drug-using period. Instead these stages refer to the concept of starting, using, and the aftermath of being exposed as a drug user by being caught or entering treatment. The *Before* stage in this paper explains the driving forces (initiation) that influence the women to turn to drugs. The *During* stage explains what the women perceive to gain from doing the drugs, even while hiding their use to appear "normal" to society. Lastly, the *After* stage explains the women's situation after being exposed as a drug user and the challenges they face, especially mothers trying to regain custody of their children.

#### A. Before Stage

There may be many reasons why women turn to drugs, including the aspect of self-medication and social pressure to fit society's high standards for "successful" women or "good" mothers. Many of the women went through emotional and psychological difficulties, such as depression and social stress, which they felt was addressed through the use of drugs. In both cases, the women began the use of drugs to avoid the stigma they faced in their current lives. The concept of "self-medication" examines coping with feelings of hopelessness and feeling different from those who seem happy in mainstream society. The social pressure to "fit in" was frequently mentioned. Specifically, idealized weight was a common reason that the women turned to methamphetamine when they felt that they were not meeting social expectations of the perfect female shape. Methamphetamine was an effective drug for losing weight.

Self-Medication. Many of the women started using drugs to help themselves recover from a break-up or separation from their significant others. For example, Linda, a 49-year-old mother, explained that drugs were what kept her sane through a divorce. She started to feel unworthy and unable to work, so she began selling and doing drugs.

"When I first started in the divorce, um, when we first separated, I was straight. I was tryin' to do right. I had the kids in church. And it got so hard and somebody was always goin' "well if you did this if you did that," and I started feelin' beneath. Uh, when I had the car wreck, I knew one way I could support my kids—I started sellin' drugs."

Methamphetamine gave Linda the energy she desired to keep up with her role as a mother, lifted her feelings of depression, and also provided a source of much needed income. Linda explained, “Yeah, you’d have to stay awake, so you used meth. With meth you feel like, uh, you’re more normal, you’re more equal to what you should be. You have energy.”

Chloe, 18 years old, had been raped a few years earlier and never reported it or received any counseling. She used drugs more often after this incident. Chloe stopped using drugs when she was in a relationship with a boyfriend who disapproved of drug use. After he left her, she felt the need to help herself emotionally and started using again.

“I think I started doing a wider variety after he left because it was like, he wasn’t around so I wasn’t afraid of getting caught. Like, he was the only person I was worried about, you know, looking down on me for doing it.”

Like Chloe and Linda, many of the women began to use drug as self-medication to make them feel better. Some had been separated from their significant others, but they revealed different reasons to use drugs to address their separation. Linda began her use of drugs as a means of financial aid to support herself as well as her children, while Chloe went back to her old habits of using the drugs she had used before meeting her boyfriend. Chloe was a previous drug user, but in this sense, the separation from her boyfriend triggered a greater variety of use.

In addition, many women faced hardships and/or difficulties in their lives that made them feel hopeless. For example, Audrey, 19 years old, explained:

“Like college is fucking hard now because I never had to study in high school and I did very well. And that was another reason why speed was such, was my drug of choice, because I would get geeked and my, like, tweak would be to do school work.”

“Speed” is a term used to mean methamphetamine. “Tweak,” to methamphetamine users, is what they like doing while high. Although Audrey was not trying to medicate herself from emotional stress, she was facing social stress in college. She used meth to try to “medicate” in the sense that it allowed her to complete her work for school. This concept of self-medication relates to the situations that the women faced that motivates them turn to drugs as a coping mechanism. Some are at a loss of what to do next with their lives after the change in social roles. Other women turn to drugs to try to improve or fix their situation and “cure” themselves—to feel normal.

Weight Loss. Many of the women also felt that there are societal pressures to being a woman that drives them to use drugs. These pressures made it stressful as well as

hard for the women to meet the standards of what they perceived are perfect women. The most difficult aspect that was a common theme brought up by the women who used methamphetamine was their weight, which they perceived as not meeting the idealized or normal weight for women of their age. Dolly, a 37-year-old mother, described her reason to use drugs, "Cause it's just hard to lose weight, you know. It's so hard to lose weight. It's not hard when you're high." Similarly, Katy, 22 years old, explained why she started using methamphetamine:

"You know, it used to be fun, hanging out with my friends, joking around, seeing idiots, but it's... you know, it was hanging out, having fun, but at the same time maintaining a weight. For girls it's a huge weight thing."

Perceived societal pressure drove these women to drug use. Some society expectations, such as maintaining an ideal weight, may not be achievable for these women without the help of drugs. Therefore, they turned to drugs as aids, leading them onto the path of further drug use. Feeling they are bound to a standard that they are not able to achieve without "help" caused them more difficulties and hardships in the next stage of their use.

### B. During Stage

Society has set norms of behavior for individuals, which translates into the standards of being "normal." Some of the women, especially the poorer women, did not feel as if they "fit in" with mainstream lifestyles and felt stigmatized. Using drugs, they were able to find an identity for themselves and, ironically, feel part of society, as discussed in the *before* stage. However, the difficulty they now faced was that as a drug user, they would still be stigmatized. Therefore, the women hid their use to avoid stigma and still meet the expectations of society. In addition, many of the women using meth felt "normal" since they now were meeting society's expectations of women's roles. They found themselves with increased energy allowing them to be more productive, as well as having increased self-esteem.

The Dilemma of Avoiding Stigma and Feeling Normal. As discussed, some of these women did not feel normal before using drugs, but felt more normal while using. Other women used drugs as an attempt to prevent themselves from being stigmatized for what they felt were shortcomings. Yet by doing this, the women risked more stigma if they were caught. The sad irony of the women using drugs to feel normal was that it actually led them to behave contrary to the norms of mainstream society. It often led to associating with other drug users, considered outside conventional society as well.

Examples of feeling normal are shown in many of their justifications for using drugs. For example, Linda explained:

“With meth you feel, like, uh, you’re more normal. You’re more equal to what you should be...Uh, you feel like you fit in, um...you feel like you’re part of the human race instead of...something that’s been thrown aside. You feel like you fit in.”

In the case of Linda, she was stigmatized by the fact that she was poor and did not meet the general expectations of mainstream middle-class society. However, by using drugs and selling them, she met the expectations of making enough money to support her family as a single mother. This aspect allowed her to feel normal rather than being tossed to the side.

Likewise using drugs to avoid potential stigma and cope with difficulties, Madeline, a 22-year-old, had been kidnapped by an abusive boyfriend when she tried to leave him. She explained her drug use eventually became normalized behavior. “And then after I was doing it for a while, it was kind of like I was doing it just to be normal,” she said, “or what I thought was normal at the time.”

Other women felt they were normal when they use drugs, but at the same time needed to avoid being stigmatized as a drug user if discovered. For example, Audrey, a 19-year-old, experienced stigma from her own parents for doing the drugs. Describing an incident with her parents, she said, “I just remember her looking at me and shaking me and saying ‘what the hell did you take?’ And my dad’s in the background screaming at me.” Audrey’s parents were upset with her, but this only caused her to hide her use, like many women who take measures to hide that aspect of their lives. As Katy, 22 years old, explained, “don’t let your family find out what you’re doing because it will hurt them the most.”

When their drug use is exposed, they are more stigmatized, even if they felt they were acting “normal” while using drugs. For example, Katy felt normal using meth, but had to hide her use to prevent the stigma, as she explained:

“My family thought I was fine. And I pulled that off for four years and no one had a...I’m sure, I’m sure somewhere my mom knew in her head I was using, but nothing pointed to it you know. I was doing everything right.”

Although Katy was still using drugs, concealing it and acting “normal” helped her avoid the stigma that would come from being an exposed drug user. Like Katy, many women started drugs to avoid feeling stigmatized, but they were now forced to hide their drug use to avoid even greater stigma. Their stories reveal how drugs seemed to help them to not only feel better but also be more productive.

Increased Productivity. Using drugs like methamphetamine, the women felt an increase in energy, which led to more productivity and better functioning in their daily lives. Linda explained,

“Well look, when I snorted it I had energy. I could get all this stuff done that, let me tell you, I cleaned my house and somebody came by one day and they said “what are you doing?” I said well I’m cleaning. I was cleaning behind my dryer. I took the metal piece off and was cleaning it (laughs).”

The increased level of energy allowed Linda to complete strenuous household chores to the standard she felt was expected of her as a mother. Similarly, Abigail, a 20-year-old, explained the utility of using meth:

“It made me like more productive. I felt like I got done stuff on it. Like I could party and hang out with my friends, and then go to work the next day, and then go to school the next day.”

These women were able to accomplish multiple tasks due to the increased energy, which also linked to avoiding stigma. They felt that by using drugs they were functioning according to the norms of society. The adverse effects of drugs were not visible yet to others, allowing these women to conceal their use.

In addition to increase in energy, meth also provided the ability to focus in order to accomplish tasks in a more productive manner. Mercedes, 34 years old, explained:

“I feel like I can hold my composure better. The thinking is much clearer...normally you can focus so clearly. Like, I used to like to do my artwork and carve when, uh, especially by hand but even with a drill and what not. But by hand it’s like I would just have an exact...I could just make the most intricate carvings. And, um, keep my mind as to different tricks to use to get different textures and different techniques. I was very, uh, just on point. Very tactful, um, and I felt that was a result of the drug.”

Mercedes also provided an example of her friend using drugs to focus:

“You know, I had one friend that would love to study with it. She would use it to study. When she had finals, she would be up for like two weeks at a time. Like usually she would buy a half-ounce of speed, you know, a couple weeks before exams.”

The ability to focus and to work on the smallest details also helped them to conceal their use of drugs from others who may stigmatize them once they find out. These

women are able to carry on with their lives and portray that themselves as “normal” and even productive to others, while feeling normal themselves only when using drugs.

Self-Esteem. Feeling normal for these women impacted their self-esteem, since their prior stigmatization from society made them feel outcast or different. Linda explained how smoking methamphetamine made her feel:

“Your self-esteem goes up with it. Um...when you start smoking it... anything, any pain anybody causes, you can kill it. You don’t have to...you can kill your emotions with it. And that’s what took over with me, with the [meth] pipe.”

The emotions that Linda referred to were the negative views the women perceived from others, but when they used meth, the emotions were subsided, ultimately increasing their self-esteem.

Isabella, a 27-year-old, had low self-esteem from having a boyfriend who physically abused her and verbally berated her for being overweight. Although her parents were upper-class and provided all her material needs, she felt ostracized by her peers. Using methamphetamine allowed her to feel better, as she explained:

“I really liked it. I just enjoyed the feeling that I got from it, is the only way I can really describe the feeling that I had. It was just I liked it. I almost felt normal, as sick as that sounds.”

The “feeling” that Isabella felt when doing drugs was reiterated by other women who said they felt they enjoyed their lives more and even felt empowered when using meth. Isabella used the phrase “as sick as that sounds,” knowing that she faced stigma within society for being a drug user. This supports the findings that many women used drugs to feel better on a personal level, despite the stigma associated with drug use.

Beyond, enjoying the feeling, there also are aspects of escape that relate to self-esteem. Mercedes, who experienced childhood rape, describes what drugs did for her. “I could escape from reality and the same time I could get high and still be on point and still be tactful,” she said. “And, in fact, it even increased that tactfulness and that energy.”

This analysis reveals that their perceived or actual social situation prevented some women from feeling valued. Beyond increasing self-esteem, when the women use drugs they enter their own social reality where things get done and they feel better. They are able to possess the abilities that they see in others that make them normal. They find value in themselves as they accomplish tasks and achieve goals that they believe society has set for them. Ironically, using drugs was a way for them to deal with feelings of being stigmatized, abused, or not valued, superficially raising their self-esteem for a time—until they are exposed as a drug user.

Being a Mother and Drug User. Being a mother and a drug user are conflicting roles. A mother is caring and supporting of the family, while a drug user is seen as an uncaring and selfish person that should be kept away from children. If a mother is a drug user, the children are believed to be exposed to a negative environment, and drug-using mothers are stigmatized in mainstream society. Several of these women expressed difficulty in maintaining their good reputation within society as a mother. This caused some to hide their drug use and only reveal what is accepted by society. Beth, a 23-year-old young mother, described how meth helped her maintain this image:

“Every bit of laundry would be done except for what they were wearing, and I couldn’t wait for them to hurry up and get undressed so I could grab those and throw them in the wash. Everybody’s socks had their initials on the bottom.”

When asked if people were suspicious, she answered:

“I don’t think so. I don’t think so. My kid’s homework was always done. Christmas time, all the school bus drivers got little baskets of cookies and homemade truffles, and I volunteered at every field trip.”

Similar to many other women, Beth was able to prevent people from suspecting she was using drugs because she maintained her reputation as being a good mother. For many, the image shattered once their drug use was discovered.

### C. After Stage

This section explores the lives of the women after they were exposed as drug users and as they faced different types of criticism. Some were struggling with the difficulty losing their children. Many faced challenges of experiencing stigma again as they were given (or assumed) the label of “former drug user.” Although many of the women were trying to improve their lives and maintain their family relationships, society still stigmatizes them because they used to use drugs (a discredited status), or because they are not as productive as they were when they hid their drug use.

Facing Criticism. Many of the women faced hardships after stopping drug use due to the criticism they felt by not meeting expectations. Lily, a 21-year-old college student, explained her relationship with her mother after she stopped using meth:

“I actually stopped before I told her. I had been off for probably three months when I told her. Because she was giving me a hard time about my living situation and she knew about me not going to classes very much. She kept saying I was fucking up my life, and I got upset about

it because I was like, you know what, I just got off of a really hard drug, and if you think that I'm such a fuck-up, well listen to this. I was on meth for a year and a half, or however long it was, and pretty much just to show her up, you know, that yeah, I have fucked up, but I'm working on it."

Lily's mother criticized her for not meeting her expectations as a daughter and a student. Others faced similar situations when they stopped using drugs to improve their lives.

Some women said they were expected to act as "normal" and productive as they were while using drugs but without the "help" that drugs gave them. For example, Katy described her interaction with her mother:

"I'm not using ice [meth] anymore. I said I've done it for four or five years, you didn't know. I hid it from you I said and I'm clean. And she's like but why have you gained so much weight. And I said, do you want me to be skinny and a crack-head, or fat and clean? Well of course I want you to be clean, but she never just dropped it at that. She would, every time I saw her I guess since April, this last April until now, just a month ago, every time I see her she won't say anything without saying something about my weight first."

Katy was criticized for not being thin as she had been when she was using methamphetamine regularly. It was difficult for Katy to stay drug-free when her mother kept bringing up the positive aspect of her drug use, in this case weight management. Their stories revealed the dilemma of using drugs to feel normal or cope with stigma, stigmatized for being a drug user, and criticized when they stopped using drugs for not meeting the standards set by society for women. It was a classic Catch-22.

Mother Roles. Many of the mothers were considered perfect role models of a mother while they were hiding their drug use. However, when the use of drugs was revealed, no matter how good of a mother they were, their children were taken away or threatened to be taken away by relatives or social services. For example, Tammy, 35-year-old mother whose children were taken from her three times, described the situation in which her drug use led to her losing her children:

"It hurt because I could remember all the good times that we used to have. I mean, you know, because basically I was, you know, I tried to be the stay-at-home mom, but, you know, I was working around the trailer park too, and all that. But just to wake up in the morning and

get up and take care of my kids, and get them off to school, and then have to sit there and wait on them to come home, you know. Me and my kids, we was inseparable. But my sister, she told me, she says, 'if you don't stop what you're doing,' she says, 'I'm going to take your kids.'"

Many of the women had expressed similar fear of losing their children or distress after losing them and having a difficult time regaining custody. Being a drug user prevents these women from fulfilling an acceptable mother role even though some effectively provided and cared for their children while using drugs. Society says they are not fit to be mothers, but it offers little help to meet the high standards set.

Although these women are drug users, they expressed care and love for their children. They tend to think of the future impact of their drug use on their children. Lisa, a 30-year-old mother, revealed that the care she felt for her unborn child while she was pregnant motivated her to stop drug use during pregnancy.

"I don't know, that's one thing I think about now, 'cause I see all my friends and stuff that are pregnant and they can't stop. But I just did, and I guess because I care more about my kid."

Likewise, Beth revealed she was concerned about the home environment of her children. The people in Beth's house, friends of her husband, were using drugs and influenced her to use drugs. After she stopped using, and in order to prevent her relapse, she encouraged her husband to ask them to leave.

"I told him, I said. Well, my thinking at that time was if they're going to do it, what the fuck, I may as well do it too, because it's going to be there. We're going to get raided, we're going to get arrested. We're going to lose the kids. If they can't pass drug tests, why the hell should I. And he went home and kicked them all out. And that was it."

Many of the women whose children were taken away, worked hard to get them back because they wanted to be mothers. Beth explained her desire to have her children back when they were taken by a relative:

"My sister-in-law that's got my middle daughter is saying, 'well, when you get better you can have her back. When you get better you can have her back.' I was going to get her back. I didn't care if I had to go to rehab for ten years, I was going to get my children back, because they deserved it."

Beth further expressed how she felt contrite for not being there for her children.

“My daughter had told me over the phone that she thought she was pregnant. Caroline was taken away from me. My youngest daughter started kindergarten. But there was a whole life that needed to be lived. (talking in very soft tearful voice). And my kids deserved the real me, not what I’d subjected them to.”

Some of the women were determined to be reunited their family, but they do not have the resources needed to be drug-free. Most of the mothers interviewed showed great compassion for their children and willingness to go to treatment to be with their family. Yet in society’s view, drug users are not “good mothers” and first they had to prove themselves. They were willing to change and attempted to get help, but many found help was not available. The formal social control agents—the police, social services and treatment providers—that coerced them to change did not provide the informal social control needed for them to remain drug-free in their own social environments.

### *SOCIAL CONTROL*

Not all women drug users have had the opportunity to receive treatment services. However, those who did found that they were not as effective as many assume them to be. Referring to the treatment services they experienced, the women explained the different aspects that could be improved. Others found religion to be helpful in remaining drug-free by living a godly life. Religion offered informal social control by providing a social group outside the drug using network, but religious belief was not for all women. The most important source of social control mentioned by the women was not treatment but instead the social environment where they lived and the social networks they had access to before, during and after using drugs. When their social environment provided positive relationships, they had enough help to stop drug use and recover over time.

#### A. Treatment Services

There are different services available for treating drug use, but many were shown to be ineffective. Even if the women were able to access drug treatment, some discovered it did not help solve the issue that they were struggling with before they started drugs or the issues they faced as a “former drug user.” Lisa explained her drug treatment experience:

“There was just, they give you, you go in Thursday, and talk about your problem, that’s all we had to do, one day a week (chuckles). So it was just inconvenience, and I guess ‘cause I had to go somewhere, but it didn’t help nothing.”

Harper, a 22-year-old, explained her experience with treatment classes in similar terms:

“I’ve been taking drug and alcohol classes for that. And I take them once a week for two and a half hours I guess...And it’s pretty good, but I still plan on smoking [marijuana] afterwards because I don’t think it’s wrong, and it helps my headaches.”

These two women considered the treatment services they attended were not very helpful, and they did not attain a drug-free life after they left the treatment services.

Many of the women found the treatment regime to be too difficult, especially drug-free programs where even prescription drugs were prohibited. On the other hand, Tammy described her experience with treatment services that provided medication in a more positive way:

“I recommend this rehab to anybody. Because when you go in, you don’t have to pay nothing. And they’ll take you and you can see, they help you with your medication. They don’t let you do no kind of, you know, hard narcotics or nothing in there, but they do help you with your, like, psychiatric drugs, you know, if you’ve got bipolar or anything. They will help you pay for them. They feed you. You go to drug classes every night. You go to, well the NA meetings. Then you go to drug classes five days a week from eight-thirty to one-thirty, and it’s a great place to be.”

Although Tammy described her program as more effective than others, she had the advantage of being offered an expensive residential treatment program with psychiatrists who treated her underlying causes of drug use. She also had substantial social support throughout her time in treatment and afterwards. Many of the other women could not access such resources, and some were told that as “former drug users” they could not use any psychotropic medication. Total abstinence was difficult to maintain.

## B. Religion

Religion seemed to play a greater role than treatment services for many of the poorer women when trying to become drug-free, especially those who did not have

access to expensive drug treatment. Dolly shared the spiritual help she felt while in drug court:

“I got the probation and all from the possession and the DUI charges. And I still got four years left on there [in drug court]. But I just moved down a level. They took me off intensive last week and I’ve got a new probation officer, and I’ve done all my community service and I’ve done everything I was supposed to do. All my programs and all my treatment and all my evaluations and therapy and everything. And it’s all helped, you know. But honestly the biggest thing is depending on God. If it wasn’t for God, I would have never made it. And I know that.”

Although Dolly thought that she has benefitted from the services that drug courts provided for her, she indicated that religion was what eventually helped her recover in her mind.

Similarly, Rachel said, “Yea, it’s about the strongest influence. My idea of God is there’s a God and He loves me a lot, and he wants what’s best for me. That’s it. And he helps me stay clean.” While her faith helped Rachel stay drug-free except for tobacco, Dolly used methamphetamine again soon after leaving drug court, but stopped using to be with her children. The act of believing in something was the motivation that allowed the women to stop using. They also believed that God was the factor that allowed them to “stay clean” and provided the help needed for them to move on with their lives.

Lisa, like many of the women, believed that the treatment centers she went to were not effective:

“I think if you found a treatment program that was like strict and some of the inpatient that you know you do get tested and stuff, and they really care, then I think that might help. But like [Generic Treatment Program] was just a joke ‘cause all you had to do was show up on Thursday, pay your ten dollars for your little meeting, talk about your problem and go home. They didn’t try to help you stop really. Cause half of our class went in high every week.”

In contrast, Tiffany, a 25-year-old, experienced the benefit of a more structured treatment environment.

“Honestly, right now it’s the fact that I’m in drug court. It’s the fact that I get random drug screens. I haven’t been tested this week but like last week I got tested three times. So it’s like very random. But going back to speed, like I graduate drug court March 3rd, Monday’s spring break. Like I’m finally going to be done with this. But I have no inten-

tion on ever using meth again. I mean I can't tell you the future, like I can't really, I'm not a psychic, but I saw what it did to me. Like when I was, when I used cocaine this last time I was in drug court and I had a positive screen and I had to go to jail for five days and I had to start over in that phase. Like, I mean I had consequences but they didn't send me back to jail for good. And I asked them, you know the officer is part of the administration in drug court and so she agreed to pull up my original felony arrest picture and looking at that, it made me never ever use it again. I looked like death walking."

The strict environment seemed to keep Tiffany on track as long as she was monitored, and it motivated her to start a new life. But it is still a coercive form of social control, which may not work for everyone, and without a structured social environment to go to when she leaves drug court, there is little hope that she will remain drug-free. As shown in her narrative, Tiffany became drug-free previously in drug court, but she relapsed after leaving because with nowhere to go, she returned to her old social environment.

### C. Social Environment

The women's accounts of ineffective services revealed that when they are not given the right resources that address the reasons they started using drugs, many turn back to the drugs that do help them cope with their personal issues. Based on their stories, a positive social environment was more important for their path to recovery than enrolling in treatment services or going to jail. After leaving treatment or jail, most of the women were back in the neighborhoods where they came from—sometimes with more drug use knowledge. "I'm not a criminal. I need help. I don't need prison," Beth explained. "Send me to prison and from what I've heard I can get more drugs in prison than anywhere else. There were drugs in the county jail."

Although the women drug users were often separated from society until they recovered, this did not help address the reasons why they used drugs. As Beth mentioned, she actually increased her use in jail and after she left. The environment of the criminal justice system was not the right social situation for women to recover.

The most influential social environment to consider outside of a treatment service is the social network of the women. When going back to their communities, many women relapsed if their friends were still doing drugs. Isabella explained:

"I mean, you've got to change your mind. But I think the biggest problem that people have, because I tried to get clean several times just because other people wanted me to, and then I would see my old friends. I mean, you have to get away from those people, even if it means pushing them away and being mean if that's what it takes. I mean, you really have to have your mind set to get away from meth."

Social support for these women was critical after they left a structured treatment environment but many of the women were exposed to people who use drugs in the neighborhoods where they return.

Some women showed how distraction away from the drugs is a key aspect that is needed in the environment for the women drug users. Harper explained how she helped a friend remain drug-free after treatment by distracting her with drug-free social activities:

“I would, I would try to distract her. Because the only thing that will really help you, because once you want to use—you want to use. And if you can get a good distraction in there, it will last for a while.”

Similarly, Tiffany described her boyfriend’s positive influence keeping her involved in activities:

“He’s just a completely positive character. Uh, he would just calm me down a lot of times. I don’t know. And he had like more of a life. Instead of like waking up every day like how am I going to get high today, his was you know, ‘let’s go do this, let’s go do that,’ you know. It was more of like activity was a way of, you know, trying not to think about it. Like he was taking me everywhere. We’d go to the islands, we’d go, you know, to the lakes, like everywhere.”

The positive social environment that these women were exposed to helped them move on with their lives. They found social support and happiness through other means than using drugs. A positive social environment after stopping drug use played a greater role in remaining drug-free than the any other solutions mentioned by the women in the study.

## V. Discussion and Conclusion

The findings in this study on women drug users show that there is not one reason for drug use and the women face several difficulties while using drugs as well as when trying to remain drug-free. Reasons for initiation included coping with the stigma of not meeting perceived social expectations, such as an idealized weight, and self-medicating to address stress or depression. The period *during* the drug use was the most difficult time in which the women tried to avoid stigma by concealing their drug use. Yet they said that they felt normal using drugs as they saw an increase in their productivity and self-esteem. The *after* stage revealed more difficulty for these women as they faced criticism on multiple levels. They were still criticized for having a “former user” label, and some women were not able to maintain the same level of productivity they showed when using meth. Another difficulty for those women who were drug users and

mothers was trying to regain custody of their children. Some of the mothers stopped their drug use after being caught to be able to be with their children. However, as a “former drug user,” society still stigmatized them and formally controlled their status of motherhood with strict rules and requirements they often could not meet. A positive social environment after leaving treatment proved to be the most helpful, such as being surrounded by supporting people who wanted them to recover.

Although a few of the women believed that the treatment services were beneficial, others found that religion was a form of social control that worked for them. Religion also provided a social environment that offered access to a new social group and support for their new way of living, but not all women could embrace religious beliefs. Others found the treatment programs were not effective because some of the people in treatment were still using drugs, or it was easy to fall back into drug use when they left with insufficient social support. Treatment with a more structured environment helped some, but this kind of treatment was not available to most of the women.

While treatment and criminal justice programs such as drug court provided formal social control, the most effective form of social control was when the women were linked to a positive social environment that gave them motivation to stay drug-free. This typically included social activities or new networks outside drug using environments. This kind of informal social control focuses on the formation of bonds and relationships outside of drug using community that helps recovering individuals maintain a drug-free lifestyle within mainstream society. The findings support other research showing that new relationships in social environments are important factors to consider when trying to help former drug users maintain drug-free lives (Boeri, Gibson & Boshears, 2014; Moos, 2007; Zschau et al. 2015).

The findings presented here add to the literature by highlighting the impact of stigma on women and how continued stigmatization hinders recovery efforts. Typically, treatment services focus on changing the individual through formal social control, but there is much less focus on changing their social environment. The findings of this in-depth study, with its focus on all stages of drug use progression, show that women need a link to mainstream society and social bonds that help motivate them to believe they can “fit in” and be normal without drugs. Simply trying to change who they are will not allow them to integrate into new social networks that provide this sense of belonging and acceptance. Continuous emphasis on their “former drug user” status, such as requiring them to go to meetings with other former users, can increase feelings of inadequacy, marginalization and stigmatization. Women drug users faced stigma in all stages of their drug use, and reducing this stigma might be the clue to successful integration into society.

*LIMITATIONS*

This is a qualitative study using a small sample of women whose primary drug of choice was methamphetamine. More research is needed to discover if other drugs, such as crack or heroin, provide the same perception of productivity and self-esteem that these women felt while using methamphetamine. Since all women in this study continued using other substances (both legal and illegal) after stopping methamphetamine, more research is needed on the potential benefits of continued short-term or long-term use of some substances compared to total abstinence. Studies with a more racially diverse sample can provide further insights on how women drug users perceive stigma and normality. The findings of the impact of stigma on women drug users might be compared to men who use drugs in future studies. Finally, many of the women in this study had experienced violent abuse, which needs to be further analyzed to better understand its impact on their drug use and recovery.

*IMPLICATIONS FOR TREATMENT AND SOCIAL SERVICES*

This study suggests that for recovering drug users, treatment services must focus on the social environment of their clients after leaving treatment. It also suggests that social services can do more to prevent women from using drugs by providing greater access to needed resources in their social environment, and intervention services for women drug users can acknowledge the social context where the women live when developing their programs. The study findings show that women need resources as well as access to new social networks. While treatment programs are the most obvious space to facilitate access to new networks for recovery efforts, social institutions, such as religious organizations, schools, and workplaces, also can provide opportunities for social activities that link women to new networks outside their own communities. Social services might facilitate participation by providing free transportation. Religious organizations can provide space for social clubs and activities for people in the community other than their own congregations or members. Workplaces can offer treatment for all levels of employees, and incentivize social activities that encourage constructive relationships. Schools can provide subsidies and transportation for children to participate in social clubs and free equipment to participate in sports teams. Incorporating an understanding of the social environment in prevention, intervention and treatment shifts the focus from individuals to the relations between individuals and social networks.

## References

- Bayer, R. (2008).** "Stigma and the Ethics of Public Health: Not Can We But Should We." *Social Science and Medicine*, 67, 463-472.
- Boeri, M. W. (2013).** *Women on ice: Methamphetamine use among suburban women*. New Brunswick, NJ: Rutgers University Press.
- Boeri, M. W. (2004).** "Hell, I'm an Addict, But I Ain't No Junkie: An Ethnographic Analysis of Aging Heroin Users." *Human Organization*, 63(2), 236-245.
- Boeri, M.W., Gibson, D. & Boshears, P. (2014).** "Conceptualizing Social Recovery: Recovery Routes of Methamphetamine Users." *Journal of Qualitative Criminal Justice & Criminology*, 2(1), 5-38.
- Boeri, M. W., Tyndall, B. D., & Woodall, D. R. (2011).** "Suburban Poverty: Barriers to Services and Injury Prevention Among Marginalized Women Who Use Methamphetamine." *Western Journal of Emergency Medicine*, 12(3), 284-292.
- Boeri, M., Gardner, M. Gerken, E., Ross, M., & Wheeler, J. (2016).** "I Don't Know What Fun Is: Examining the Intersection of Social Capital, Social Networks, and Social Recovery." *Drugs and Alcohol Today*, 16(1), 1-11.
- Boyd, Susan C. (1999).** *Mothers and Illicit Drugs: Transcending the Myths*. Toronto: University of Toronto Press.
- Campbell, Nancy. (2000).** *Using Women: Gender, Drug Policy, and Social Justice*. Psychology Press.
- Charmaz, K. (2006).** *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Thousand Oaks, California: SAGE Publications Inc.
- Chaudoir, S. R., Earnshaw, V. A., & Andel, S. (2013).** "'Discredited' Versus 'Discreditable': Understanding How Shared and Unique Stigma Mechanisms Affect Psychological and Physical Health Disparities." *Basic and applied social psychology*, 35(1), 75-87.
- Connera, K. O. & Rosen, D. (2008).** "'You're Nothing but a Junkie': Multiple Experiences of Stigma in an Aging Methadone Maintenance Population." *Journal of Social Work Practice in the Addictions*, 8(2), 244-264.
- Copes, H., Hochstetler, A., & Williams, J. P. (2008).** "We Weren't Like No Regular Dope Fiends." *Social Problems*, 55(2), 254-270.
- Dluzen, D. E. & Liu, B. (2008).** "Gender Differences in Methamphetamine Use and Responses: A Review." *Gender Medicine*, 5(1), 24-35.
- Glaser B. G. & Strauss, A. (1967).** *Discovery of Grounded Theory*. New York: Aldine.

- Goffman, E. (1963).** *Stigma: Notes on the Management of Spoiled Identity*. New York, New York: Simon & Schuster, Inc.
- Goffman, E. (1959).** *The Presentation of Self in Everyday Life*. New York, New York: Bantam Doubleday Dell Publishing Group, Inc.
- Hall, E. A., Baldwin, D. M., & Prendergast, M. L. (2001).** "Women on Parole: Barriers to Success After Substance Abuse Treatment." *Human Organization*, 60(3), 225-233.
- Hetrick, E. S., & Martin, A. D. (1987).** "Developmental Issues and Their Resolution for Gay and Lesbian Adolescents." *Journal of Homosexuality*, 14, 25-43.
- Jiménez, J., Puig, M., Sala, A.C., Ramos, J. C., Castro, E., Morales, M., & Zorrilla, C. (2011).** "Felt Stigma in Injection Drug Users and Sex Workers: Focus Group Research with HIV-risk Populations in Puerto Rico." *Qualitative Research in Psychology*, 8(1), 26-39.
- Laub, J. H., & Sampson, R. J. (2003).** *Shared beginnings, Divergent Lives: Delinquent Boys to Age 70*. Cambridge, MA: Harvard University Press.
- Lende, D. H., Leonard, T., Sterk, C. E., & Elifson, K. (2007).** "Functional Methamphetamine Use: The Insider's Perspective." *Addiction Research & Theory*, 15(5), 465-477.
- Link, B.G., & Phelan, J.C. (2001).** "Conceptualizing Stigma." *Annual Review of Sociology*, 27, 363-385.
- Luck, P. A., Elifson, K. W., & Sterk, C. E. (2004).** "Female Drug Users and the Welfare System: A Qualitative Exploration." *Drugs: Education, Prevention and Policy*, 11(2), 113-128.
- Lyons, T., Shannon, K., Pierre, L., Small, W., Krüsi, A., & Kerr, T. (2015).** "A Qualitative Study of Transgender Individuals' Experiences in Residential Addiction Treatment Settings: Stigma and Inclusivity." *Substance Abuse Treatment, Prevention, and Policy*, 10(1), 1.
- Maher, L. & Hudson, S. L. (2007).** "Women in the Drug Economy: A Meta-synthesis of the Qualitative Literature." *Journal of Drug Issues*, 805-826.
- Malterud, K. (2001).** "Qualitative Research: Standards, Challenges, and Guidelines." *The Lancet*, 358, 483-88.
- McKenna, K. Y., & Bargh, J. A. (1998).** "Coming Out in the Age of the Internet: 'Identity Demarginalization' Through Virtual Group Participation." *Journal of Personality and Social Psychology*, 75(3), 681.
- Moos, R.H. (2007).** "Theory-Based Active Ingredients of Effective Treatments for Substance Use Disorders." *Drug and Alcohol Dependency*, 88(2-3), 109-121.
- Neale, J., Nettleton, S., & Pickering, L. (2011).** "Recovery From Problem Drug Use: What Can We Learn From the Sociologist Erving Goffman?" *Drugs: Education, Prevention and Policy*, 18(1), 3-9.

- Nettleton, S., Neale, J., & Pickering, L. (2012).** "I Just Want to be Normal": An Analysis of Discourses of Normality Among Recovering Heroin Users." *Health*, 17(2), 174-190.
- Nichter, M. et al. (2006).** "Gendered Dimensions of Smoking Among College Students." *Journal of Adolescent Research*, 21(3), 215-243.
- Olsen, A. (2015).** "Punishing Parents: Child Removal in the Context of Drug Use." *Drug and Alcohol Review*, 34(1), 27-30.
- Pachankis, J. E. (2007).** "The Psychological Implications of Concealing a Stigma: a Cognitive-Affective-Behavioral Model." *Psychological Bulletin*, 133(2), 328.
- Palamar, J.J. (2012).** "A Pilot Study Examining Perceived Rejection and Secrecy in Relation to Illicit Drug Use and Associated Stigma." *Drug and Alcohol Review*, 31, 573-579.
- Palamar, J. J., Halkitis, P. N., & Kiang, M. V. (2013).** "Perceived Public Stigma and Stigmatization in Explaining Lifetime Illicit Drug Use Among Emerging Adults." *Addiction Research & Theory*, 21(6), 516-525.
- Quinn, D. M., & Chaudoir, S. R. (2009).** "Living with a Concealable Stigmatized Identity: the Impact of Anticipated Stigma, Centrality, Salience, and Cultural Stigma on Psychological Distress and Health." *Journal of Personality and Social Psychology*, 97(4), 634.
- Radcliffe, P. (2011).** "Motherhood, Pregnancy, and the Negotiation of Identity: The Moral Career of Drug Treatment." *Social Science & Medicine*, 72, 984-991.
- Radcliffe, P. & Stevens, A. (2008).** "Are Drug Treatment Services Only for 'Thieving Junkie Scumbags'? Drug Users and the Management of Stigmatized Identities." *Social Science & Medicine*, 67, 1065-1073.
- Roberts, D. E. (1991).** "Punishing Drug Addicts Who Have Babies: Women of Color, Equality, and the Right of Privacy." *Faculty Scholarship*. 1370.
- Sered, S. S., & Norton-Hawk, M. (2014).** *Can't Catch a Break: Gender, Jail, Drugs, and the Limits of Personal Responsibility*. Oakland, CA, University of California Press.
- Stengel, C. (2014).** "The Risk of Being 'Too Honest': Drug Use, Stigma and Pregnancy." *Health, Risk & Society*, 16(1), 36-50.
- Stone, R. (2015).** "Pregnant Women and Substance Use: Fear, Stigma, and Barriers to Care." *Health & Justice*, 3(1), 1-15.
- Strauss, A. & Corbin, J. (1998).** *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. 2nd edition. Thousand Oaks, CA: Sage

- Walter, T., Ford, A., Templeton, L., Valentine, C., & Velleman, R. (2015).** "Compassion or Stigma? How Adults Bereaved by Alcohol or Drugs Experience Services." *Health & Social Care in the Community*. doi: 10.1111/hsc.12273.
- Watters, J. & Biernacki, P. (1989).** "Targeted Sampling: Options for the Study of Hidden Populations." *Social Problems* 36, 416-30.
- White, W. L. (2007).** "Addiction Recovery: Its Definition and Conceptual Boundaries." *Journal of Substance Abuse Treatment*, 33(3), 229-241.
- White, W. L. (2002).** *Long-Term Strategies to Reduce the Stigma Attached to Addiction, Treatment, and Recovery Within the City of Philadelphia*. Philadelphia: Department of Behavioral Health and Mental Retardation Services.
- Woodall, D., & Boeri, M. (2013).** "When You Got Friends in Low Places, You Stay Low: Social Networks and Access to Resources for Female Methamphetamine Users in Low-Income Suburban Communities." *Journal of Drug Issues*, 44(3), 321- 339.
- Zschau, T., Collins, C., Lee, H., & Hatch, D. L. (2015).** "The Hidden Challenge: Limited Recovery Capital of Drug Court Participants' Support Networks." *Journal of Applied Social Science*, 10(1), 22-43.

## YAY OR NEIGH: IS THE HORSE RACING MARKET EFFICIENT?

By Jillian Raia\*

*Billions of dollars are bet on horse racing each year. People are attracted to the thrill of the race, the possibility of winning big, and the conviction that there is in fact a method to picking the winners. This study will focus on the latter, questioning whether or not the horse betting market exemplifies underlying principles of the efficient market hypothesis (EMH); in short, whether it is possible to 'beat the market.' This study begins with an analysis of the EMH from a financial markets and economics perspective, then applies these theories to the horse racing market, and concludes with evidence from a survey I created to poll horse racing sentiments from a variety of bettors. The research demonstrates the horse racing market is efficient in both the weak and semi-strong forms. However, the market seems highly inefficient in the strong form. In addition to the literature, the results of my survey reveal important distinctions between the casual bettors and handicappers. Interestingly enough, the competition between these two betting classes helps pull the market towards efficiency in the long-term indicating we as bettors might be running a race of our own.*

Keywords: Horse racing; Efficient markets hypothesis; Handicapping; Above average returns; Betting; Weak form; Semi-strong form; Strong form.

### I. Introduction

On a single June day in 2014, more than 100,000 people were there in person to watch California Chrome attempt to make Triple Crown history (Kissell, 2014), wagering nearly eight million dollars on the outcome (McGrath, 2015). At home and at bars, more than 20 million people viewed and bet another \$83 million. If there was ever any doubt, these numbers demonstrate that thoroughbred horse racing is an extremely popular event and gambling outlet in the United States. But is it possible to consistently make money? My research focuses on whether or not the horse race betting market is efficient or inefficient, and the betting behaviors affecting this market.

The first part of this paper will outline the most salient features of the efficient market hypothesis (EMH) and its important elements from a business perspective. Since 1970, economists have debated the EMH, which was developed by the American economist Eugene Fama. Underlying his thesis is the assumption that the stock market is efficient because it is impossible to consistently 'beat the market' by earning

---

\* Email: jraia2494@gmail.com. I would like to thank my advisor, Professor Christopher Beneke for his endless support, wisdom, and encouragement.

above average profits. In other words, the market typically reflects all relevant information without a noticeable lag; therefore, preventing investors from steadily turning a significant return. Fama has drawn significant support from other economists, as well as opposition. Those who oppose the EMH note how some investors are able to exploit market information to their advantage. In addition, these scholars address several anomalies, which investors can benefit from. I will also take into consideration insights drawn from a particular set of critics known as behavioral economists who recognize that people often act in economically irrational ways.

After describing the EMH and its weaknesses, I will address some of its key features in light of the horse racing market. Fama's theory typically refers to the stock market; however, Durham, Hertz, and Martin (2005) suggest using sports betting markets as an empirical setting. I will show that many EMH economic concepts may be applicable to the horse betting market, but only with significant qualifications. The horse racing market contains anomalies and fosters an atmosphere where decisions are often affected by bettor's behaviors. I will address both sides of this argument by applying the weak, semi-strong, and strong forms of the EMH to horse racing in an attempt to shed light on the complexities of this market. Finally, I will consider what the casual bettor can take away from such research. After all, people have been traveling to the racetrack for hundreds of years, with the confidence of ending the day with more money to their name. The question I confront is this: Is there a viable strategy for 'beating this market' or should horse racing fans simply take pleasure in the thrill of the race and the wagers they place on them?

## II. EMH in the Stock Market

The efficient market hypothesis has been an intensely debated topic in economics for decades and is likely to remain so for some time. In 1965 Eugene Fama began conducting research on the use of technical analysis and fundamental analysis to predict the movement of stock prices. In his paper, "Random Walks in Stock Market Prices" (1965), Fama begins by describing technical and fundamental analysis theories. He explains that technical analysts believe past prices will be reflected in future prices, encouraging them to search for patterns that will indicate stock market history is repeating itself. Fundamental analysts focus their attention on intrinsic values. All securities have an intrinsic value at any point in time, also known as an equilibrium point, which is determined by several fundamental factors. In studying these fundamental elements, such as quality of management, industry outlook, level of competition, etc., one should be able to determine if a stock's intrinsic value is above or below its actual price. An investor would then take a position with regard to the difference between both of these values. Both technical and fundamental analysts rely on using past and present data to predict future values in an effort to gain a significant return.

Throughout his paper, Fama explores the phenomenon of "random walks" within the stock market, which challenge the principles underlying technical and fundamental

analysis. A random walk indicates that the movements of a security are unpredictable in regard to future changes because the value of said security is likely to unpredictably rise and fall (Mishkin, 2012). In describing the concept of a random walk Fama begins by defining an efficient market as “a market where there are large numbers of rational profit-maximizers actively competing, with each trying to predict future market values of individual securities, and where important current information is almost freely available to all participants” (Fama, 1965). Furthermore, when a market is considered efficient “competition will cause the full effects of new information on intrinsic value to be reflected ‘instantaneously’ in actual prices” (Fama, 1965). In other words, stock prices will accurately reflect all relevant information, leaving technical and fundamental analysts at a loss. Malkiel (2003) takes Fama’s assumptions a level deeper explaining that, “tomorrow’s price change will reflect only tomorrow’s news and will be independent of the price changes today. Since news is, by definition, unpredictable, the resulting price changes must also be unpredictable.” Fama concludes in his 1965 paper that the stock market is in essence a series of random walks that are “no more predictable than the path of a series of cumulated random numbers.” In essence, it is impossible to forecast prices that result from an unpredictable path.

Fama’s 1965 random walks paper led to his most notable work, “Efficient Capital Markets: A Review of Theory and Empirical Work,” where he develops the EMH, making random walks the underlying basis. In plain terms, it is impossible for one to beat the market if it is efficient and follows an unpredictable path. Fama (1970) breaks down his tests of efficiency into three different levels: weak form, semi-strong form, and strong form. The weak test of efficiency is concerned with historical prices, essentially indicating that all current prices should reflect information from past price movements. The semi-strong form assumes that current stock prices will reflect all public information both past and present. The strong form asserts that market prices accurately echo all information both publically held, as well as privately held. In broader terms, the weak test of efficiency is designed to discredit the use of technical analysis, while the semi-strong test argues that both technical analysis and fundamental analysis are meaningless. The strong form of market efficiency is intended to caution that not even privileged information can lead to consistent, above-average returns on the market.

There are thousands of empirical studies supporting the EMH. For instance, Mishkin (2012) discusses the *Wall Street Journal*’s “Investment Dashboard,” which shows the inconsistent return analysts redeem from market predictions. A random selection of stocks are chosen by throwing darts at a financial page of the newspaper. This random selection is then compared to buy and sell recommendations from a group of advisors. Interestingly enough, the advisors and the dartboard exchanged victories indicating the advisor’s investment strategies were unable to achieve more significant returns. Degutis and Novickyte (2014) correspondingly address such evidence, claiming “if actively managed portfolios fail to outperform passive portfolios, it is then not profitable to collect market information, and the market is efficient.” This support highlights Fama’s idea that the market follows a random walk or unpredictability.

While Fama gives an influential argument for the EMH offering much explanation and evidence, there are also several scholars who oppose the EMH. Mishkin (2012) dedicates an entire appendix to discrediting the EMH. Anomalies, which are occurrences that deviate from the standard or norm, provide the most conflicting evidence. Mishkin (2012) first addresses the small-firm anomaly, which demonstrates through many empirical studies that “small firms have earned abnormally high returns over long periods of time.” In fact, since 1926 small firm stocks in the U.S. have produced rates of return over one percentage point higher than larger firms (Keim, 1983, as quoted in Malkiel, 2003). However, Malkiel (2003) argues that higher returns from smaller companies are not able to reveal predictable patterns. Furthermore, despite Malkiel’s (2003) note of how profitable small firms have been since 1926; recently the level of returns small firms have generated has declined since the 1980s.

Another anomaly commonly discussed is the January effect, which posits that stock prices tend to rise abnormally from December to January. This is usually due to the holiday season; however, the importance behind the anomaly is the predictability. Investors can adequately predict that stocks will deviate from “random walks” during this January period, challenging the concept of market efficiency. Although these anomalies might be present, Malkiel (2003) believes “there is no way in which investors can reliably exploit any anomalies or patterns that might exist.” This author does not argue the existence of such anomalies, but instead focuses on their inconsequential nature. Since market efficiency implies investors are unable to earn *consistently* above-average returns, scholars such as Malkiel would argue anomalies are not a viable exception to the EMH.

Market overreaction also contradicts the belief that it is impossible to beat the market. Mishkin (2012) turns economist’s attention towards excessive volatility, meaning fluctuations of stock values can in fact be much greater than the fluctuations in their fundamental values. This source mentions a study done by Robert Shiller to support this idea. Shiller (1981) conducted research regarding fluctuations in the S&P 500 to determine if the variations in this index could be explained through the dividends of the comprised stocks. He was unable to fully determine the cause of movement in the index indicating that “stock market prices appear to be driven by factors other than fundamentals” (Shiller 1981, as quoted in Mishkin, 2012). In other words, market prices are not entirely representative of publicly held information, as the EMH states.

Contradictions of the EMH are abundant, yet one in particular stands out—behavioral finance. Behavioral finance directly challenges the concept of efficient market through the idea that people do not act rationally, but instead are driven by a number of emotions. Degutis and Novickyte (2014) discuss investor’s irrationality and the term ‘gambler’s fallacy,’ which can be defined as the tendency for people to distort the probabilities of an event occurring based on past outcomes. Singh (2010) also addresses the credibility of behavioral finance indicating it challenges two major assumptions of the EMH, “1) that the majority of investors make rational decisions based on available information; and 2) that the market price is always right.” The argument is

that EMH cannot be valid, for people's investment decisions are influenced by more than just rational behavior backed by public information. That being said, the market price is flawed due to the presence of irrational investors. As Hersh Shefrin argues in Singh (2010), scholars of behavioral economics "believe investor psychology can drive market prices and fundamental value very far apart." Singh (2010) continues to address some of these human biases and emotions, such as representativeness, confirmation bias, quantitative anchors, and moral anchors. Representativeness suggests that people may notice a pattern and assume future performance will follow without any true analysis or understanding of the pattern. Confirmation bias explains the habit of seeking information that will support existing beliefs without regards to the quality of the information. Quantitative anchors encompasses the idea that people are fearful of the unknown. Human tendency is to make a choice with whatever information is at hand even if it is unreliable or irrelevant. In other words, people overweigh the past due to the immense uncertainty of the future. Lastly, moral anchors are stories that help justify ones actions. Hearing a personal story about someone buying a lottery ticket and winning a million, might encourage one to buy a ticket of their own. These scholars mentioned, as well as several others, believe behavioral finance serves as a direct challenge to the premises of the EMH. Douglas (1995) goes so far as to claim the EMH only "works by coincidence." He believes in order for the theory to hold true, one must make "domain assumptions" in regards to the market that are unrealistic, such as assuming all investors will act rationally. Interestingly enough, many of these irrational human tendencies serve as major players in the horse racing market, disturbing efficiency.

Many scholars and investors argue anomalies, volatility, and human behavior cause inefficiencies in the market that can be exploited; however, advocates of the EMH cling to the word *consistency*. They contend that such inefficiencies are only present for a limited amount of time, but then become obsolete when learned by the society. When an investor discovers a way to exploit the market it will be successful as long as it remains isolated. Malkiel (2003) explains it well stating, "the more potentially profitable a discoverable pattern is, the less likely it is to survive." He uses the January effect as a prime example of what once was a profitable anomaly that resulted in self-destruction. According to the January effect, stocks tend to be extremely profitable during the first five days of January. Investors have tried for years to take advantage of such "market inefficiencies." However, this became a common strategy and in order to profit investors had to buy their stocks earlier in December and sell them earlier in January, until this method became useless. Malkiel (2003) also describes this knowledge sharing phenomenon through the story of \$100 bill. Malkiel (2003) tells a famous story of a student and a finance professor who were walking down the street. When the student sees a \$100 bill on the ground, he bends down to pick it up. The finance professor stops him and says, "Don't bother—if it were really a \$100 bill, it wouldn't be there" (Malkiel, 2003). Once again, the idea is that profit opportunities do not last very long. In essence, scholars addressing the short lifespan of lucrative investment

strategies are indirectly contending that markets will reflect all publically available information, making them efficient at least to a semi-strong extent.

There is a constant battle between market efficiency advocates and their opposition. The two sides continuously debate about what seems to be a circular form of reasoning. The interesting paradox of the EMH is mentioned in Degutis and Novickyte (2014) in reference to Grossman (1976), where the academic proposed “the more investors believe in market efficiency, the less efficient the market becomes.” Grossman goes on to explain that if people think the market is efficient, hence follows a “random walk,” they will not probe for information as intensely. In response, inefficiencies in the market will increase. Once again, think about Malkiel’s (2003) \$100 bill story. If people stop searching for \$100 bills, then there will be more of them on the ground. Such a paradox can be seen in horse race betting. If people stop probing for information to make educated bets then the chance of making a profit on the winning horse will increase.

As evident, economics literature regarding the EMH is abundant, yet there remains no solid conclusion on whether or not the market is efficient or inefficient. Both sides offer empirical studies and analyses to support their theories. Advocates of the EMH would argue that there lacks substantial evidence of investor’s ability to earn consistently above-average profits, which validates Fama’s theory. Many supporters will not argue the presence of anomalies; however, they will discredit their power. As Shiller (2003) expressed in his studies, “the [pricing] anomalies that had been discovered [in the 1970s] might be considered at worst small departures from the fundamental truth of market efficiency, but if most of the volatility in the stock market was unexplained, it would call into question the basic underpinnings of the entire efficient market theory” (as quoted in Singh, 2010). In other words, short-lived, minimal profitability obtained from anomalies is not enough to disprove the EMH. On the contrary, the opposition would contend the market does not “fully reflect” all relevant information, for if it did then people would not be able to exploit it. Furthermore, the EMH fails to mention the behavioral effects people have on the market. In essence, claiming the market is efficient is endorsing the idea that “market analysts and fund managers work entirely by luck” (Douglas, 1995). While both sides present valid support for their beliefs, the argument remains at a standstill.

### **III. EMH in the Horse Racing Market**

Significant similarities between the stock market and the horse racing market allow us to analyze the presence of market efficiency at the race track, and determine how profitable one can be betting on the races. Ottaviani (2009), Asch (1982), Hausch (1981), and Thaler (1988) all emphasize how truly analogous these markets are. Asch et al (1982) outline some comparisons such as the uncertainty of future earnings, the large number of participants, access to a broad range of information, and the presence of professional opinions. Hausch et al (1981) are in agreement with this thought as they

explain such similarities would lead to a breadth of fundamental and technical strategies designed to beat the market. This is demonstrated by the abundance of research and published material regarding strategies for thoroughbred horse racing. Thaler and Ziemba (1988) take this association deeper in stating wagering markets are better environments for testing efficient market theories due to the “well-defined determination point [of each bet] at which its value becomes certain.” In other words, the horse racing market has a distinguished end. The bets in this market have a definite lifecycle that provide quick and repeated feedback. As a result, Thaler and Ziemba (1988) indicate the horse racing market has a greater chance of being efficient for bettors because they can learn from each race. On the contrary, “a stock is infinitely lived, its value today depends both on the present value of future cash flows and on the price someone will pay for the security tomorrow” (Thaler and Ziemba, 1988). In horse racing, a thoroughbred’s future performance has no bearing on a present race, in effect making only historical information relevant. Ottaviani and Sorensen (2009) agree and contend, “given the frequent observation of the realized outcomes and the absence of bookmakers, pari-mutuel betting markets offer an ideal test bed for theories of information aggregation and market efficiency.” Although their analyses and conclusions may differ, each of these scholars concludes the horse race betting market provides a useful setting for a discussion of the EMH.

The horse race betting market is considered pari-mutuel meaning all of the money wagered on a race goes into one pool. Based on the outcome of said race, this money pool is distributed among the winners in relation to the amount each person wagered. In essence, pari-mutuel bettors are betting against each other, as the losers go home with nothing and the winners allocate the loser’s money. Before this pool of money is distributed a “track take” is removed as compensation for the track, its operators, and the government. Typically “track take” is between 15 and 20 percent. There is also an element of the payouts called “breakage,” meaning all returns per dollar bet are rounded down to the nearest five or ten cents (Thaler and Ziemba, 1988). Together, both of these payout elements are considered transaction costs. We can use a simple formula to summarize the return potential in a pari-mutuel betting system:

*Total betting pool – track take + breakage = amount to be paid out to winning wagers*

In horse racing, the market lasts for about 25 minutes. This is the amount of time bettors have to select the horse or horses they believe will finish “in the money,” implying first, second, or third place. In horse racing terminology, that is win, place, or show, respectively. Payouts are made on bets that correctly backed horses in the money. Racegoers consist of a diverse group of people, but in terms of betting strategy, they tend to fall into one of three categories. First, there are the handicappers, who are considered the professionals in the field of horse betting. They devote much of their time trying to exploit inefficiencies in the market through both fundamental and techni-

cal strategies. Then there are bettors who attempt to make an educated bet by reading the top picks, briefly analyzing the odds, and using the few tips they have read or heard about prior to arriving at the racetrack. Lastly, there are those who are simply there for the experience and are compelled to bet just because they can. This category does not employ any betting strategy, but tends to focus instead on the thoroughbreds' names and the colors the jockeys are wearing.

### *WEAK FORM*

Considering the comparability of the horse racing market to the stock market, the same highly debated question of market efficiency applies to the horse racing arena. It is practical to begin with a discussion of weak form efficiency, which implies current market prices accurately represent past performance. In the horse racing market, evidence of the efficiency in the weak form would demonstrate the odds projected for each horse at the start of the race mirror each horse's actual chance of winning, and the bettor's inability to earn above average return solely using this information. As stated by Ottaviani and Sorensen (2009), "the efficient market hypothesis asserts that the fraction of money wagered by the market on an outcome is an unbiased estimate of the outcome's empirical frequency." In essence, a weak form of efficiency in the horse racing market implies that people's bets will reflect relevant information and should accurately depict each horse's empirical probability of winning. At the beginning of each race day, a professional determines what are known as the "morning line odds." This professional analyzes a series of horse racing data in an attempt to predict how the public will bet on each race. However, these odds will change frequently up until the start of the race, due to the pari-mutuel nature of the sport. There are twenty-four cycles or points at which the tote board will display the updated odds (Asch et al, 1982). According to the EMH, people react to all publicly available information and this is exhibited in the odds; therefore, these "market values" should be accurate representations of a horse's potential.

In their study, Asch et al (1982) examined 792 races from Atlantic City Race Course, which included a total of 5805 horses. They found that the objective odds, which are the actual measurable results, have a high correlation to the subjective odds, which are those determined by the public.<sup>1</sup> As Asch et al (1982) describe it, "the subjective odds of the betting public do a good job overall predicting the true objective probabilities." Ottaviani and Sorensen (2009) and Snyder (1978) all consider this a weak test of the EMH because the analysis focuses on whether or not the odds provided at the track can be used to earn above average profits. To employ a useful analogy, security prices are to the stock market as subjective odds are to the horse racing market. According to the EMH, the security prices of a stock will reflect all past prices and investors will not be able to consistently beat the market using them. In horse racing,

---

1. Objective odds- a horse's actual chance of winning; Subjective odds- the probability a horse will win determined by the public. These odds are influenced by the public's wagers.

the subjective odds will also reflect all historical information and consequently bettors will not be able to earn above average profits using them. This concept is known as price discovery, which is when prices eventually settle at values that represent a horse's true winning probability. The strong correlation between actual odds and pari-mutuel odds, described by Asch et al (1982), support the belief that the horse racing market is weakly efficient.

In terms of return, an efficient horse racing market should demonstrate a series of random walks, thus no abnormal profit potential for investors in the long-term. Asch et al (1982) explain that in an efficient market "rates of return for betting horses with shorts odds ought to be the same as the results of betting horses with intermediate odds [and] with long odds." Likewise, Snyder (1978) claims if the horse racing market was efficient, "then the expected rate of return for all types of bets would be identical and simply equal to the (negative) track take." Snyder (1978) and Asch et al. (1982) complement each other's analysis. Over the long term there will not be significant returns in an efficient market. Instead, the average bettor should lose about 18% of her investments, or an amount equivalent to track take. This is because favorites will win more often, but produce smaller payouts, while longshots rarely win, but offer larger payouts. The market will settle on equilibrium--so whether you bet on all favorites, all longshots, or a mixture the results should follow a random walk and hence lead to market efficiency.

Nevertheless, Asch et al (1982) notice an interesting anomaly in their data, commonly referred to as the favorite-longshot bias. In the stock market, anomalies indicate inefficiencies. As depicted in Figure 1, favorites win 36.1 percent of the time, but the crowd only predicts them winning 32.5 percent of the time. In contrast, the least favorite horses (or longshots) win 1.7 percent of the time, yet the bettors predict them winning 2.5 percent of the time. In essence, the favorites are often underbet while the longshots are often overbet, which is the basis of the favorite-longshot bias. In the case of the favorite, odds will be higher than warranted and the payout better. The opposite holds true for the longshots. The takeaway is that favorites typically are better bets than longshots, for they have a greater expected return. At the extremes, subjective probabilities continuously deviate from fundamental probabilities revealed in the market, exhibiting signs of inefficiency.

FIGURE 1  
Taken From Asch (1982), p. 191

Table 1  
Subjective and objective probabilities of winning in 729 Atlantic City (NJ) races in 1978 (total number of horses = 5805).

Favorites <sup>a</sup> (1)	No. of races <sup>b</sup> (2)	Obj. prob. <sup>c</sup> (3)	Subj. prob. (4)	(Subj. prob. – obj. prob.)/ st. error of obj. prob. <sup>d</sup> (5)
1st	729	0.361	0.325	–2.119 <sup>e</sup>
2nd	729	0.218	0.205	–0.903
3rd	729	0.170	0.145	–1.972 <sup>e</sup>
4th	724	0.115	0.104	–0.961
5th	692	0.071	0.072	0.074
6th	598	0.050	0.048	–0.279
7th	431	0.030	0.034	0.480
8th	289	0.017	0.025	1.096
9th	165	0.006	0.018	2.095 <sup>e</sup>

<sup>a</sup>Lowest odds horses.

<sup>b</sup>The number of races declines because many races have only a small number of entrants. It should be noted that there are numerous races in which there is a tie for which horse is the first favorite, or second favorite, etc. The pool of first favorites was taken to consist of all horses with the lowest odds, including ties, and similarly for the other positions.

<sup>c</sup>Note that these probabilities are the probabilities for the *i*th favorite conditional on there being an *i*th favorite in a particular race. Hence they need not sum to unity.

<sup>d</sup>The standard errors were computed by taking the objective probabilities as the 'true' probabilities and assuming a binomial process. Thus the standard error is  $[p(1-p)/n]^{1/2}$  [see Ali (1977)], where *p* is the objective probability and *n* the number of races.

<sup>e</sup>Significant at the 0.05 level.

The returns demonstrated through the favorite-longshot bias also contradict the market efficiency belief. For example, Figure 2 shows that betting on a favorite with odds between zero and two will lead to a -13.7 percent rate of return, which is better than the average of -18 percent. The rate of return decreases to -63.7 percent when betting on longshots with odds above twenty-five (Asch et al, 1982). This indicates that it is more profitable to bet on the favorites than betting on the longshots. Furthermore, betting on favorites yields a return greater than the efficiency rate of -18 percent. A different experiment conducted by Snowberg and Wolfers (2010), exhibits similar results with a sample of 5,610,580 races in the United States from 1992 to 2001. The study shows that betting on horses with odds greater than 100 to 1 will yield a -61 percent rate of return, random wagering will yield -23 percent, and betting the favorite in every race yields a -5.5 percent rate of return (Snowberg and Wolfers, 2010). Snyder (1978)

also examines this common phenomenon by summarizing six studies regarding the favorite-longshot bias—Fabricant (1965), Griffith (1949), McGlothlin (1956), Seligman (1975), and Weitzman (1965). All six of these experiments also illustrate horses with short odds are more profitable than those with long odds.

## FIGURE 2

Taken From Asch (1982), p. 192

Table 2

Rates of return from bets on horses with different odds levels for all races and for late races (races 8 and 9).

Odds level <i>O</i> (1)	Rates of return	
	All races (2)	Late races (3)
$0 \leq 2$	-0.1366	-0.0428
$2 < 0 \leq 3.5$	-0.3177	-0.3210
$3.5 < 0 \leq 5$	-0.1758	-0.0288
$5 < 0 \leq 8$	-0.2242	-0.5238
$8 < 0 \leq 14$	-0.1602	-0.1698
$14 < 0 \leq 25$	-0.3255	-0.3618
$25 < 0$	-0.6372	-0.6858

At first glance, evidence of such return patterns might appear to signal market inefficiency in the weak form; however, the conclusive factor here is profitability. Bettors are not able to overcome track take through exploiting this inefficiency in the long-term (Ashe et al, 1982). Each of these studies indicates betting on favorites offers a greater return than betting on longshots; however, it is critical to note that none of these strategies offer positive returns in the long run. Attempting to exploit the favorite-longshot bias in every race will not render profitability for the bettor. This concludes the horse racing market is efficient in the weak form.

The presence of the favorite-longshot bias as an inconsequential market inefficiency is evident, but the reason for such a bias enhances complexity. Just as investor's behavior affect their decisions in the stock market, human emotion affects people at the racetrack. However, the question becomes whether people are actively seeking risk in betting longshots or if they do not understand the probabilities at hand. Snowberg and Wolfers (2010), consider these two interesting theories of risk-love and misperceptions in an attempt to explain the favorite-longshot bias. Risk-love indicates "the favorite-longshot bias can be fully rationalized by a standard rational-expectations expected utility model" (Snowberg and Wolfers, 2010). This theory is consistent with neoclassical economics, which studies supply and demand assuming humans act in a rational manner. However, economic evolution has introduced behavioral economics, which recognizes people do not always act logically. As a result, the misperceptions theory

assumes people are not risk-loving, but in fact do not truly understand the probabilities they are betting on.

Snowberg and Wolfers (2010) utilized a data set stretching from 1992 to 2001 of 6.4 million horse race starts in the United States to test which theory is responsible for the favorite-longshot bias, using exotic bets to infer preferences and perceptions regarding win bets. According to their results, the inefficiencies caused by the favorite-longshot bias are better explained through the misperceptions theory, which assumes “the low rates of return to betting longshots are rationalized by bettors who bet as though horses with tiny probabilities of winning actually have moderate probabilities of winning” (Snowberg and Wolfers, 2010). In other words, most bettors at the track do not completely understand the true probabilities they are betting on. Instead, these objective odds are distorted by human behavior so that a 30 to 1 shot seems only like a 15 to 1 shot when the bettor begins envisioning herself winning a large sum of money, or being correct in selecting the longshot underdog. Figure 3 presents their findings, which indicate that the misperceptions model mimics actual data more closely in exacta, quinella, and trifecta bets.<sup>2</sup> Taking a closer look at the exacta outcomes shows that the misperceptions model was 57.9% closer to the actual results of each race compared to only 42.1% of the risk-love model (Panel A of Figure 3). Snowberg and Wolfers (2010) contend this extensive study demonstrates the favorite-longshot bias can be explained by a variety of nuances, but the most persuasive explanation are bettors’ misperceptions. People tend to underbet favorites and overbet longshots because they do not truly understand the probabilities illustrated on the tote boards.

Another intriguing concept that demonstrates the effect of bettor’s behavior in the market is the prevailing presence of the favorite-longshot bias at the end of the day. Thaler and Ziemba (1988), Asch et al. (1982), Snyder (1978), Hausch et al. (1981), and other scholars indicate throughout their research that the favorite-longshot bias is even more distinct during the last race of the day. Bettors overbet longshots even more during the last race in an effort to recoup the losses they have suffered. This pattern provides additional evidence of bettor’s irrational behavior, which can be driven by misperceptions. Thaler and Ziemba (1988) address several emotional factors that alter betting decisions, the first being “bettors might overestimate the chances that the longshots will win.” Such misperceptions are even stronger at the end of the day when bettors are trying to break even. Thaler and Ziemba (1988) also note the simple thrill of holding a longshot ticket, the bragging rights one attains when winning on a 30 to 1 shot, and bets placed for unrelated reasons all have a hand in the favorite-longshot bias.

---

2. Exacta- a bet in which you predict the first and second place finisher in that order; Trifecta- a bet in which you predict the top three finishers; Quinella- a bet in which you predict the top two finishers in no particular order.

FIGURE 3  
Taken From Snowberg and Wolfers (2010)

Table 1: Mean Error Based Tests of Risk-Love versus Misperceptions Model

<i>Test:</i>	(1) Absolute Error:  Prediction – Actual	(2) Absolute % Error: $\frac{ Prediction - Actual }{Actual}$	(3) Which Prediction is Closer to Actual? (%)
Panel A: Exacta Bets (n=197,551)			
Risk-Love Model	0.0139	34.3%	42.1%
Misperceptions Model	0.0125	28.0%	57.9%
Risk-Love Error – Misperceptions Error	0.00137 (.00002)	6.3% (.1%)	
Panel B: Quinella Bets (n=70,169)			
Risk-Love Model	0.0274	39.0%	46.0%
Misperceptions Model	0.0258	36.3%	54.0%
Risk-Love Error – Misperceptions Error	0.00155 (.00003)	2.7% (.2%)	
Panel C: Trifecta Bets (n=137,756)			
Risk-Love Model	0.00796	100%	28.9%
Misperceptions Model	0.00532	57.4%	71.1%
Risk-Love Error – Misperceptions Error	0.00264 (.00001)	42.9% (.2%)	

*Notes:* Standard errors in parenthesis. Predictions and actual outcomes are measured in the price of a contract that pays \$1 if the event occurs, zero otherwise.

A comprehensive analysis of the weak form of market efficiency in relation to horse racing, demonstrates that bettors do not act rationally in part due to misperceptions and create an inefficiency in the market known as the favorite-longshot bias. However, this does not provide enough evidence to deem the market inefficient. The variation between actual odds and the bettor's final pari-mutuel odds is not large enough to overcome the track take of 18 percent. Therefore, bettors are not able to earn above average profits in the long-term through exploiting the favorite-longshot bias and the market is considered efficient in the weak form.

#### *SEMI-STRONG FORM*

Upon deeming the market weakly efficient, we now move to a semi-strong test of the horse racing market. If the horse racing market is efficient in the semi-strong form, we would expect to see educated bettors, known as handicappers, earning above average returns using a variety of publically available market inputs. According to Fama (1970), the semi-strong form is concerned with whether or not past and present information can be used to predict prices. If bettors are able to use various elements of

public information to predict a horse's odds of winning and consistently earn a profit then the market is considered inefficient in the semi-strong form. Handicappers spend countless hours analyzing an array of horse data such as, previous record, trainer, jockey, race distance track condition, and more, in order to gain a competitive advantage. According to Thaler and Ziemba (1988), these serious bettors use fundamental strategies, which arise from publically available information. Handicappers attempt to use a horse's past performance to predict future performance, while at the same time searching for variations between a horse's intrinsic value and actual value. In other words, these educated bettors will only place bets on pari-mutuel odds they believe to be beneficially misrepresented.

At first consideration, one might be surprised at how difficult it can be to pick a winner at the racetrack. It is a race, so doesn't it make sense to pick the horse that has the best recorded time? If it were that easy we would all be making millions. Horse racing is intriguing because no two races are ever the same. A horse's times are affected by location, track condition, distance, the competition, and an immense number of additional factors. For instance, Andrew Beyer (1993) affirms "Spectacular Bid came into the 1979 Kentucky Derby winning three 11/8-mile prep races on three different tracks labeled 'fast,' in times of 1:50, 1:482/5, and 1:484/5" (p.11). In reality, a horse's performance is difficult to predict because there are several inconsistent factors that make up their objective probabilities.

Andrew Beyer, a columnist for the Washington Post is a horse racing expert who devotes his time and effort to creating what are known as Beyer's speed figures to help him gain a competitive advantage in the horse racing market. He uses various elements of the *Daily Racing Form* (DRF), a public source of information, to assign each horse a speed figure. While the algorithms used to calculate these speed figures are complex and still remain unknown to the general public, their application is not. The horse with the highest speed figure in a given race will most likely be victorious. According to Beyer, his speed figures "simply combine a speed rating and track variant and express how fast a horse ran in past races" (p.15). A speed rating translates previous race times into comparable numbers, while the track variant measures the speed of the racing surface (Beyer, 1993). Speed figures are listed as whole integers and typically range from 0 to 130. Figure 4 exhibits a page taken from the *DRF* for the eighth race at Aqueduct Racetrack. Note the speed figures assigned to SayHey CJ are boxed off in black. Throughout his book Beyer praises the success of speed figures, but cautions that one must be able to correctly interpret the numbers and correlate them with other factors, such as pace, track bias, and trips. One must be able to distinguish those occasions when speed figures are reliable from those in which a slower horse might prevail. After all, there are countless number of elements that impact thoroughbred horse racing outcomes.

Beyer has devoted four books to the study and explanation of his speed figures, but admits such strategy can be summed up in just a few short paragraphs (Beyer, 1993). A bettor should begin by analyzing the speed figures listed in the *DRF* for the horses in a given race by following these steps:

- (1) Take a look at each horse's most recent speed figure and determine which one has the highest number for this "becomes the frame of reference for the entire race" (p.30).
- (2) Next, determine if the distance and track conditions this horse is racing in today are similar to those that were present in the race in which he earned this noteworthy speed figure. This will be a subtle indication if the horse has the ability to earn a similar speed figure today.
- (3) Subsequently, it is important to analyze the competition. Horses that do not have any speed figures similar to the horse you have chosen as your reference can be disregarded. By contrast, those who have similar speed figures or a speed figure better than the one pinpointed must be examined further.

Beyer explains simplistically that "a superior betting situation arises when a horse's last two figures are both larger than everybody else's last two" (p.31). This is considered a "double fig" and has a large chance of winning, with triple figs and omnifigs being even more superior. The preceding analysis is the most rudimentary use of Beyer speed figures, which was once fairly successful. However, like any good investment methodology, Beyer speed figures eventually began to lose their competitive advantage as they became more popular and well-known to the crowd.

FIGURE 4  
Daily Racing Form

# 8

## Aqueduct

1 MILE (1:32<sup>3</sup>) CLAIMING, Purse \$14,000 For Four Year Olds And Upward Which Have Never Won Three Races Or Four Year Olds, Weight, 123 lbs. Non-winners of two races at a mile or over since February 8 Allowed 3 lbs. Such a race since then Allowed 5 lbs. Claiming Price \$7,500

## Clim 7500B



**1 SayHey CJ**  
 Own: Blue dawg Stable  
 2-1 Blue, White N, Blue Cap  
 RUSSELL S (69 9 5 11 .13) 2009: (878 104 .12)

**B, g, 7 (Mar)**  
 Sire: Malibu Moon(A, P, Ind) \$40,000  
 Dam: Endeite (Thirty Eight Paces)  
 BR: South Norwood Breeders  
 TR: Salen Bobby A(22 1 3 6 05) 2009: 213 27 .13

15Feb10 -8Lrl fst 7f :22:45:108<sup>8</sup> 1:20<sup>3</sup> 3: Gen GrgeH-G2  
 16Feb10 -3Aqu fst 1<sup>1</sup> :23:47:111<sup>1</sup> 1:22<sup>3</sup> 4:1 EvngAttr65k  
 2Jan10 -8Lrl fst 1 :24:47:143<sup>1</sup> 1:38 4: NatvDancer57k  
 18Dec09 -8Lrl fst 1<sup>1</sup> :24:48:143<sup>1</sup> 1:44<sup>1</sup> 3: OC 50K/65-N  
 28Nov09 -8Aqu fst 6f :21:44:56<sup>1</sup> 1:09 3: FallHwh-G3  
 24Nov09 -9Lrl sly 6f :21:44:56<sup>1</sup> 1:09 3: DeFrncsM-G1  
 15Aug09 Deaווille (F) gd 1<sup>1</sup>, T RH 2:04 4: Prix Gontaul-Biron-G3  
 Racing Post Rating: 105 SIK 113590

27Jul09 -9Crl gd 6f :22:45:57<sup>1</sup> 1:09 3: Chesapeake48k  
 4Jy09 -8Del fst 6f :21:44:57<sup>1</sup> 1:10 4: OC 100k/c

20Jun09 -8CT fst 7f :23:47:111<sup>1</sup> 1:24<sup>3</sup> 3: WildNWond100k  
 16May09 -10Pm fst 6f :23:45:57<sup>1</sup> 1:09 3: MidSprnth-G3  
 18Apr09 -10CT fst 1<sup>1</sup> :47:11:14:136<sup>1</sup> 1:49<sup>1</sup> 4: CTClassic615k

**WORKS:** • Feb-L38<sup>1</sup> fst :36 B 1/5 Dec12L11 fst :1:42 B 1/1  
 TRAINER : Spmtroute (15 .13 81.13) Dir1 (216 .12 \$1.03) Routes(76 .11 \$1.30) Aw (22 .14 \$0.82)

Life	37	14	6	4	3892,524	110	D, Fst	27	8	4	4	\$524,030	110
2010	3	0	0	2	\$27,450	94	Wet (370)	8	6	1	0	\$351,994	110
2009	11	4	0	0	\$166,920	101	Synth	0	0	0	0	\$0	-
Lrl	18	9	3	3	\$447,030	110	Turf (276)	2	0	1	0	\$16,500	82
Greenspring117 <sup>1</sup> ;							DST (356)	5	2	2		\$117,780	105
Undrstrmint117 <sup>1</sup> ;													
WPostono116 <sup>2</sup> ;													
StormPly116 <sup>2</sup> ;													
Firm hold inside early													
77-09													
Encaustic122 3 3/4;													
SayHey CJ 120 <sup>2</sup> ;													
Rail turn, split foes													
77-28													
Understatement20 <sup>2</sup> ;													
J desRevenge 120 <sup>2</sup> ;													
Rail duel, came back													
96-28													
SayHey CJ170 <sup>2</sup> ;													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													
93-12													
Cherokee-Country129 <sup>1</sup> ;													
Greenspring125 <sup>1</sup> ;													
Hatfield123no Late gain wide													

### A. Loss of Informational Advantage

Similar to the way one loses an informational advantage in the stock market, the same occurs in the horse racing market, supporting the idea that it is impossible to continuously earn above average profits. In *Beyer on Speed*, Beyer describes how he began using these figures for his personal gain, but soon they lost some significance. Beyer took pride in the horse racing analysis he did, explaining, "It was as if I were interpreting the data in the *Daily Racing Form* with a code that nobody else possessed" (p.2). However, soon Beyer had company. By 1980 the concept of handicapping became popular with dedicated horse bettors employing their own methods to create a systematic set of numbers on which they could base their picks. Ultimately, the *DRF* began searching for handicapped speed figures they could list in their own forms. On April 25th, 1992 the *DRF* began incorporating Beyer's speed figures into their publications (Beyer, 1993). As Beyer explained, "the numbers that [he] had once guarded jealously became a part of the record of every horse at every recognized track in the United States and Canada" (p.12). Beyer had encountered a problem; his speed figures were losing their informational advantage.

This idea of obsolete information recognizes that there may be inefficiencies in the market that can be exploited, but once this competitive advantage becomes public knowledge it loses its effectiveness. The pari-mutuel structure of horse racing fosters competition among bettors and directs profits to those who have discovered strategies hidden in the information provided. According to Beyer (1993), "the nature of the sport ensures it cannot have a Rosetta Stone," for if there was an instruction manual for deciphering winning horses everyone would be using it (p.32). This would enhance market efficiency and drive down the profits for the informed bettors. Beyer explains that when he first started using speed figures he was able to collect earnings on horses with odds of 50 to 1 because few bettors saw the horses' potential. However, Beyer's speed figures are now published in the *DRF*, which drives down their value. Beyer notes that his speed figures are no longer very advantageous standing alone, but coupled with other key information, such as pace and track bias, these numbers still provide bettors with a meaningful strategy.

### B. Pace

Pace is a critical factor to consider when analyzing speed figures because it helps decipher how a horse earned its figures. In essence, pace seeks to legitimize a horse's speed figures. Beyer (1993) observes, "when a horse is able to take a clear early lead, running at a moderate or slow pace, he will almost always earn the best speed figure of which he is capable" (p.56). This might seem counter-intuitive at first, for we might think of ourselves running a race. Most people will run faster if they have someone running with them that is able to challenge their pace. However, this is not the case in horse racing. According to Beyer (1993), "the optimal way for a horse to run is to

get clear, to set a pace well within his capacity, to conserve energy and release it in the final stages of a race” (p.56-57). That being said, a horse might demonstrate an exceptional speed figure because he was never challenged or jostled throughout his run. On the other hand, a horse might exhibit a slower speed figure than the rest of his performances because he had another horse in his path early on and was forced to run at his competition’s pace.

Beyer (1993) elaborates by explaining how horses have different preferences and may be able to thrive due to the pace that is set during the race. There are slow-pace races where the horse who takes the lead sets a slow and steady pace. This way he can save his energy and give his all in the final stretch of the race to produce a victory. On the other end of the spectrum are fast-paced races, in which an unusually fast preliminary pace will destroy the performance of speed horses who are supposed to win. Some horses are entered into a race with the sole purpose of exhausting the superior horses. Aside from taking the lead at the start of the race, “a horse’s best possible tactical position is to be sitting third, by himself, behind a pair of dueling speed horses” (Beyer, 1993). In this case, the horses battling head-to-head in the first and second positions will diminish each other’s chances of victory. For instance, the Kentucky Derby is an extremely popular race, which is known to attract sprinters. Beyer (1993) notes the 1982 Kentucky Derby as a prime example, as “the horses who were running 19-18-16-14 after the first half mile of the Derby finished 1-2-3-4...” (p.65). Consequently, analyzing a horse’s speed figures in concurrence with their preferred paces will provide a bettor with a useful indication of which horses are most likely to win.

### C. Track Biases

Racetracks are considered biased similar to the way MLB baseball parks can be considered “home-run friendly” or not. Each thoroughbred track has its own set of unique advantages and disadvantages. According to Beyer (1993):

“No factor can render speed figures so meaningless. The greatest of horses can be defeated by bias; ordinary horses can be turned into champions by a bias....All serious bettors therefore pay keen attention to this aspect of the game” (p.71).

Tracks can demonstrate biases due to their speed or a favoring rail position. Beyer (1993) indicates four important distinctions that can be determined by reading the footnotes in the *DRF*: tracks with a good rail (GR), tracks with a bad rail (BR), speed-favoring tracks (S), and closer-favoring tracks (C) (p. 72). Beyer warns against imposing a bias when one does not exist; however, being aware of them will help determine the validity of a horse’s speed figures.

Overall, Beyer (1993) addresses pace, track bias, and several other factors to indicate how significant these elements are in horse racing. While Beyer’s speed figures

alone might have offered a competitive advantage in the past; they no longer do so. These figures must be used in conjunction with other fundamentals for handicappers to benefit from them and for the horse racing market to demonstrate signs of inefficiency.

#### D. Analysis of Handicapping

Handicappers are similar to the educated investor trying to earn a profit in the stock market through careful analysis, research, and perhaps a new strategy that remains temporarily unknown to the public. As a result, this class of bettors, along with their strategies of speed figures, pace, and track bias are used to evaluate the horse racing market in the semi-strong form. It is reasonable to assume an analysis of Beyer speed figures and complimentary information would give handicappers a considerable advantage in the horse race betting market. According to a study done by Rosenbloom (2003), a model using Beyer speed figures is better than the efficient market model. Rosenbloom (2003) begins his paper acknowledging the prevailing opinion that horse racing betting provides an efficient market. Hence the foundation of a winning strategy is one that forecasts better odds than determined by the public's pari-mutuel odds. Rosenbloom (2003) claims that using Beyer speed figures, indicative of a horse's past performance, will indicate a horse's present performance. He uses a statistical test known as a SPRT test to compare the Beyer speed figure model to the pari-mutuel model shaped by the bettors at the track using a data set comprised of 1056 horses racing in the fall of 1996 or the spring of 1997 (Rosenbloom, 2003). In his conclusions, Rosenbloom claims "the result of the SPRT like test establishes that the [Beyer speed figures] model provides more accurate probabilities than the efficient market theory." In other words, Beyer speed figures are a useful indication of a horse's chance of winning and more often than not forecast better odds than the public. When Beyer speed figures predict a horse's actual chance of winning as 3 to 1, but the public project this horse at 10 to 1 through pari-mutuel betting, this would be a wise bet. Although Rosenbloom's (2003) statistical study yields positive results surrounding handicapped bets, the sample size was small. Additional research is needed to conclude semi-strong market inefficiency.

In addition to Beyer speed figures there are several other sources of information expert bettors use in handicapping. For instance, post position is extremely important because where a horse begins a race may be indication of how well he can obtain a front running position out of the gate.<sup>3</sup> Every race track releases data regarding which post positions have the best winning probabilities, and many handicappers incorporate this information into their picks. Jockey, trainer, and owner are additional critical pieces of information. While it is the horse that runs, a jockey needs to know how to properly ride his horse in order to achieve its best performance. Different trainers have different strategies and approaches to preparing race horses to compete. Similarly, owners will only purchase a horse that matches his repertoire. Another important factor for some handicappers is a horse's latest workouts. Workouts are popular sources of information

---

3. Post position is the position a horse leaves the starting gate from. (i.e. post position 1 is closest to the rail)

because they provide the most current measure of a horse's speed and ability. Figure 4 shows workout information listed in the *DRF* boxed off at the bottom. The material includes: date of the workout, track, distance, track condition, time, if the horse was being pushed (distinguished by a B for breezing or H for handily), and rank compared to the other horses that worked out (*DRF*). A bullet indicates a workout was the fastest of the day. Additionally, many handicappers believe layoffs are critical. A horse is considered "laid off" if he has not raced in 45 days. It is common opinion that if this is a horse's first race since being "laid off" he will not perform well. Lastly, it is noted if a horse is on any particular medication, such as Lasix. In Figure 4, the bolded "L" at the top of the *DRF* indicates Lasix, which are believed to improve a horse's performance. Overall, bettors have an abundance of information available to them. It is important to distinguish between what is important and what is not in order to have a valuable handicapping strategy.

A majority of scholars would indicate the horse racing market is efficient in the semi-strong form because handicappers cannot consistently earn above average profits. By contrast, many handicappers would contend there are several inefficiencies in the market one can exploit. If handicappers did not feel this way they probably would not be funneling millions of dollars into horse betting. However, Beyer makes an interesting statement in *Beyer on Speed* (1993), contending "even if a premise is fairly logical, no simple, mechanical way of playing the horses can succeed in the long run" (p.32). In essence, Beyer who is a handicapper himself is indicating the horse racing market is efficient in the long-term. And yet, billions of dollars are bet on horse races each year. The truth is, the horse racing market is efficient in the semi-strong form, as well as the weak form; however, due to a highly educated class of bettors and an abundance of information it might appear otherwise. The market is efficient if a bettor places a wager in every race. Yet, if a bettor tactfully selects his participation in this market he can serve as evidence of the market's ability to be inefficient. For instance, in an article titled "Triple Crown Teachings," Andrew Beyer counsels, "Knowing which horse is likely to win doesn't make you money; you need to know which horse is mispriced" (Mauboussin and Callahan, 2015). Beyer is indicating that there is an important distinction between a horse's value and price. When a market is efficient, there is little to no variation between a security's intrinsic value and its price in the market. The same can be said for the horse racing market when comparing a horse's true probability of winning and a horse's payout determined by the public. The market remains efficient until anomalies and bettor behavior drive value and price in opposite directions. This is where educated investors and handicappers can take advantage of their respective markets.

Horses often diverge from their intrinsic values at the end of the racing day due to uninformed bettors allowing their emotions get the best of them. Figure 2 illustrates interesting data regarding the last race of the day. Consistently betting the favorite in the last race equates to a negative four percent rate of return. This is significantly

higher than the result of betting the favorite in all the other races (14%) due to the intensified emotions of the general public. It is common for a majority of casual bettors to wager on a longshot during the last race of the day in an effort to go home a winner and recoup previous losses. In essence the favorite-longshot bias is amplified here, as the chart demonstrates. Many of the favorites are overpriced, which presents an advantageous inefficiency in the market if one is able to recognize it. The favorite should have a high chance of winning, but due public sentiment his odds will be longer, thus producing a larger rate of return. Handicappers look for these opportunities and several others in order to be profitable.

### *STRONG FORM*

The final test of market efficiency is the strong form, which examines whether or not insiders such as jockey, trainers, certain handicappers, and others with access to restricted information have the ability to beat the market. Evidence of these insiders earning above average returns using their privileged information would distinguish the horse racing market as inefficient in the strong form. Snyder (1978) seeks to test this strong form using 846 races from Arlington Park, Chicago. He uses five different sources to represent the expert opinion: Arlington Park's track handicapper, the *DRF*, and three Chicago newspapers (Snyder, 1978). Three of these publications employ full-time professional handicappers- the track handicapper, the *DRF*, and the *Sun Times*. Two do not- the *Daily News* and the *Chicago Tribune*. Snyder's (1978) study focuses on comparing the expert predictions to those of the general public to see if there is any correlation. The results indicate that the expert's predictions demonstrate a greater bias than the general public. In other words, the experts seem to accentuate the favorite-longshot bias by "placing too few horses in the lowest odds-classes and this results in their high rates of return...[and] placing too many horses in the [highest] odds-classes, causing the rates of return to be highly negative" (Snyder, 1978). Such an outcome is surprising, for these five handicapping sources are expected to have more relevant knowledge than the general public. However, "[Snyder's] interviews with the experts revealed that they do not attempt to predict each horse's actual winning chances, rather they attempt to estimate the odds the public will create through its pari-mutuel betting" (Snyder, 1978). In other words, when graphed the predictions determined by Arlington Park's handicapper, the *DRF*, and these three Chicago newspapers should closely mimic the public's rates of return.

It is interesting that the odds quoted in the *DRF* and newspapers as morning line odds, which bettors are able to use as a guide, are sending them in a direction of greater bias than our own human nature already imposes. As a result, the distinction between the two types of handicappers is crucial. There are those who determine the morning line odds and those who forecast a horse's actual probability of winning. Predicting the morning line odds is in reality forecasting how the betting public will bet, or where the money will come in. Snyder's (1978) study focuses on this breed of handicappers.

According to these experts interviewed by Snyder, “they did not quote the favorites at lower (and more realistic) odds or the longer shots at higher (and more realistic) odds because they didn’t want to influence the public’s betting behavior by ‘pointing a finger’ at those horses more likely to win or lose” (Snyder, 1978). In effect, morning line handicappers do not want to make picks too obvious for the betting public. This expert bias is accentuated to the extreme in the last race of the day. Snyder (1978) demonstrates the difference between the public’s rates of return and the five expert rates of return. Results show that the track handicapper differed from the public by 6.7 percent in the first eight races, but three times more in the last race. Similar results can be seen with the *DRF*, *Sun Times*, and *Daily News*. Overall, Snyder (1978) reaches the conclusion that his strong test of the EMH indicates the horse racing market is efficient. The morning line odds determined by expert handicappers cannot be used to earn an above average profit.

Losey and Talbott (1980) challenge Snyder’s study arguing the five expert handicappers he used to form his conclusions do not provide an adequate test of the strong form of market efficiency. These handicapping predictions are made using publically available information and the “odds are published prior to races with the express purpose of providing handicapping aid to the betting public” (Losey and Talbott, 1980). As a result, these handicappers cannot be considered insiders. They are not seeking to take advantage of restricted information. Furthermore, the goal of these handicappers as stated in Snyder (1978) is to predict the betting behavior of the general public, not the actual odds of a horse winning. In opposition, Losey and Talbott (1980) consider the test employed by Snyder (1978) to be instead a test of semi-strong market efficiency. If we know that the morning line odds, which are expertly determined, differ from the public’s pari-mutuel odds bettors should be able to exploit this discrepancy. For example, if Horse A has morning line odds of 3 to 1, but the public bets move this horse to 8 to 1 this should signal profit potential. Losey and Talbott (1980) note that “if the market imperfectly incorporates this information, this will allow an astute bettor an opportunity to make above average returns.” In essence, if the public is not taking the morning line odds into consideration then the market is not fully reflecting all available information, and thus is inefficient. However, this study demonstrates the opposite is true showing that overlay ratios ranging from 1.01+ yield increasingly negative returns as the ratio becomes greater (Losey and Talbot, 1980). These scholars conclude that the market appears to be inefficient at first glance due to the presence of handicapper overlays; however, the inability to turn a profit using such differences indicates market efficiency.

Although both the Snyder (1978) study and the Losey and Talbott (1980) paper offer interesting insight to tests of efficiency in the horse racing market, their focus on morning line odds seems immaterial. The goal of these handicapping experts is to predict the betting actions of the public. Yet, it is evident people do not act rationally and public misperception is common at the track. To compare predictions of irrational bettors (morning line odds) to how these uncertain bettors actually wagered does not

seem effective. Instead, an overlay between the betting public and the forecasted probability of a horse winning determined by a legitimate insider would be useful.

Thought-provoking data regarding odds within the last minutes to post (MTP) indicate there may be subtle evidence of insiders' ability to exploit market inefficiencies. Asch et al. (1982) specify the significance of the time informed bettors place their wagers in the presence of a pari-mutuel betting system. Because the odds of each horse are determined by the bettors themselves, those who have any inside information will want to wait until the last minute to place their bets so they do not signal the winning horse to the general public. According to the results, "the marginal odds of the late bettors appear to be at least as good as and perhaps better than the final odds in predicting order of finish" (Asch et al, 1982). This study indicates that informed bettors tend to pick the winners, which would indicate some level of inefficiency in the market. However, due to the small data set of only 729 races additional support is needed. Gramm and McKinney (2009), produce a similar study regarding the timing of "late money" using a sample size of 1644 US races from 64 different tracks. Figure 5 illustrates the results from this analysis (Gramm and McKinney, 2009). Looking at the win column, the favorite wins 36.1 percent of the time. The informed bettors' pari-mutuel odds, which compute to 36.3 percent come extremely close to the objective odds of 36.1 percent. A similar pattern can be seen in the place and show pools. These results suggest that those with insider information, who place their bet at the last possible moment, do in fact hold knowledge of valuable information. The profit potential of such information remains uncertain. However, Gramm and McKinney (2009) point out that informed bettors pull the market closer to efficiency. As exemplified in Figure 5, the post-time subjective probability was 32.6 percent; however, after incorporating the informed wagers the final subjective probability rose to 34.2 percent. Although this information is being integrated into the market odds; there may still be room for exploitation because a slight bias still exists. The favorite will win on average 36.1 percent of the time, but the market still forecasts a probability of only 34.2 percent indicating the subjective odds will be higher.

Overall, the evidence on behalf of strong form market efficiency is limited; and therefore no meaningful conclusions can be made. The betting behavior of insiders is hard to obtain and public dissemination of such information would negate any inefficiencies – and therefore any advantages – discovered. If insider information is valuable enough to earn above average profits in horse race betting, the users of such data will do everything they can to make sure it stays secret. Upon entering the public domain, such "insider information" loses its value and its informational advantage. However, it is reasonable to believe the insiders, who work closest with a horse will know when he is going to run well today or when he is unhealthy, impacting his performance.

FIGURE 5  
Taken from Gramm and McKinney (2009)

**Table 1. The timing of bets**

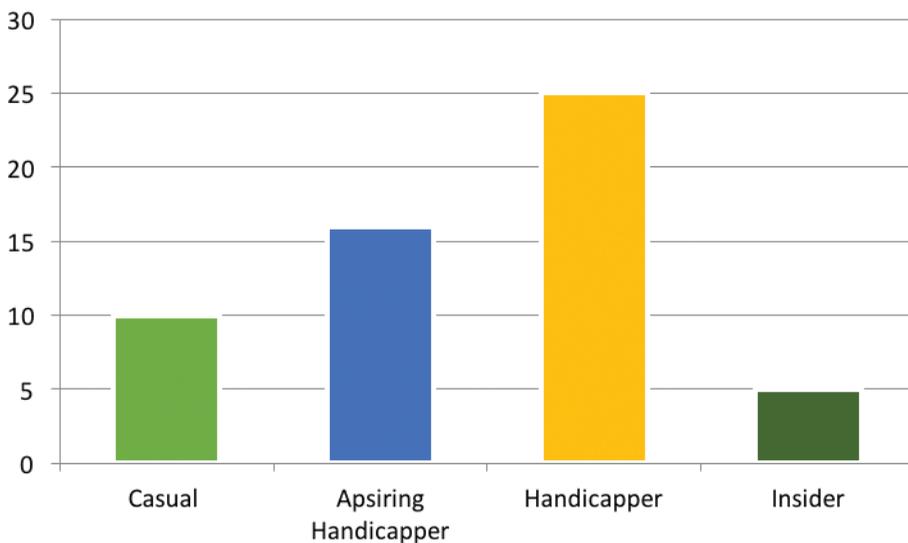
Favourite position	Win (%)			Place (%)			Show (%)					
	Objective probability	Final subjective probability	Post-time subjective probability	Late subjective probability	Objective probability	Final subjective probability	Post-time subjective probability	Late subjective probability	Objective probability	Final subjective probability	Post-time subjective probability	Late subjective probability
1	<b>36.1</b>	34.2	<b>32.6</b>	<b>36.3</b>	57.2	<b>54.7</b>	<b>51.5</b>	58.3	70.1	69.5	<b>66.0</b>	71.9
2	22.7	21.1	<b>20.6</b>	21.9	43.4	<b>40.8</b>	<b>39.6</b>	43.2	60.3	<b>55.7</b>	<b>53.6</b>	<b>62.9</b>
3	14.0	15.2	15.2	15.3	31.8	32.5	32.1	33.3	48.0	47.1	46.1	49.7
4	8.5	<b>10.7</b>	<b>11.0</b>	<b>10.3</b>	21.5	<b>24.5</b>	<b>25.0</b>	<b>23.9</b>	37.9	38.6	38.5	38.1
5	7.6	7.4	7.9	6.8	18.6	18.1	19.0	16.8	31.9	30.7	31.9	<b>27.5</b>
6	5.6	5.2	5.7	<b>4.5</b>	13.1	13.1	14.3	<b>11.4</b>	23.1	24.3	<b>25.9</b>	21.6
7	3.7	3.7	4.1	3.0	9.7	9.6	10.7	8.1	18.5	19.2	<b>20.9</b>	16.7
8	2.6	2.7	3.1	2.2	6.2	7.2	<b>8.2</b>	5.8	12.0	<b>15.1</b>	<b>16.9</b>	12.3
9-14	0.9	<b>1.8</b>	<b>2.1</b>	1.3	3.7	<b>4.8</b>	<b>5.6</b>	3.6	7.7	<b>11.0</b>	<b>12.6</b>	8.5
$\sum_{i=2}^n$		28.02	49.30	20.25		23.97	67.99	20.00		49.56	132.11	29.27
%Pool		100.0%	61.4%	38.6%		100.0%	62.5%	37.5%		100.0%	64.6%	35.4%

Note: Bold type indicates statistically different from objective probability.

## IV. Survey Methodology and Evidence

Most of the academic evidence indicates the horse racing market is efficient, yet there are several other factors to consider affecting market efficiency such as, emotion, bettor knowledge, and experience. While scholars contend it is impossible to consistently beat the horse racing market, thousands of people travel to the track and approach the betting windows to place wagers. In an effort to discover what drives people to do so, I developed a survey to gauge the horse racing sentiments of various bettors. My survey, which was distributed to family and friends, was comprised of fourteen questions that were a mixture of multiple choice and open response (see Appendix). I could offer my own opinion on what persuades me to wager, but I wanted to obtain feedback from a reasonable number of diverse bettors. I was advised that it would be better to restrict my survey to family, friends, and friends of friends in order to obtain reliable results. Although I needed an adequate number of responses, reliability was my main focus upon distribution. I created and distributed my survey through Qualtrics.com, which is a well-renowned marketing research website that offers customers the ability to create their own confidential surveys. Their website provided me with a link to my survey, which I emailed to family and friends encouraging them to participate.

**FIGURE 6**  
Betting Class Breakdown (N = 56)



As seen in Figure 6, the data are comprised of 56 responses representing a wide range of bettors, from horse aficionados to those who go to the track for entertainment. Casual bettors are those who travel to the track for the social experience and bet simply because they can. This group makes up 18% of the people polled. I defined aspiring

handicappers as those who place data-driven bets and try to implement strategy behind their wagers. Handicappers, on the other hand, devote countless hours to analyzing horse racing data in an effort to develop a reliable betting strategy. 28% of respondents were aspiring handicappers, while 45% were professional handicappers. Lastly, I received 5 responses (9%) from insiders, which included trainers, jockeys, owners, and anyone else who has access to restricted information. This well distributed data set allowed me to analyze the underlying sentiments that foster competition among bettors and contribute to the market's efficiency.

Certain trends exhibited in the results regarding amount bet, MTP, and frequency appear to represent their betting classes well. For instance, of the 30 respondents that were either insiders or handicappers, 73% of them bet more than \$20,000 on horse racing each year. For many this may seem like a large sum to spend on gambling; however, handicappers and insiders would disagree. These horse aficionados view thoroughbred horse racing as a means of generating additional income, a game, and even an art. The handicappers and insiders represent serious bettors who believe they can exploit inefficiencies in the market either through intense analysis of past and present data or privileged knowledge. They are not testing their luck, but instead they are testing their skills at handicapping. Another thought-provoking characteristic of these bettors is that all 30 of them, with the exception of four, placed their bets with less than nine minutes-to-post (MTP). In my research, I found that informed bettors have a tendency to wait until the last minute to place their bets because they do not want to tip off the general public. The data in my survey correspond to this assertion made by Asch et al (1982). Furthermore, since these bettor seem to be wagering large sums of money it is reasonable for them believe their picks will alter the tote board's probabilities, thus warning the general public of a valuable pick. The frequency with which these respondents bet is also significant because it indirectly exemplifies the attitudes and objectives of this class of serious bettors. For instance, 35 people responded "other" when asked how frequently they place bets, indicating they bet more than once each week. One would assume, the more serious the bettor, the more often he would wager. This was exactly the case in my survey. Of the 35 bettors who responded "other" 71% of them were handicappers or insiders. This betting class spends a lot of their time researching horses and their past performances; hence they try to bet as often as possible given the exploitation opportunities they discover. As we can see, the actions of this sophisticated class of bettors relate directly to their passion for thoroughbred horse wagering.

In order to demonstrate how people react to situations differently, I presented a hypothetical betting condition to those polled and asked them to respond accordingly. I expected to see the causal bettors displaying more signs of emotional betting in comparison to the serious bettors. However, this trend did not follow suit. Hypothetically, it was the eighth and final race of the day and the bettor had lost all of the earlier races. How would each bettor wager in this final race? I gave respondents three choices: don't bet at all, wager less, wager an amount consistent with your other bets today, or wager more. The literature demonstrated that people tend to increase their wagers at the end

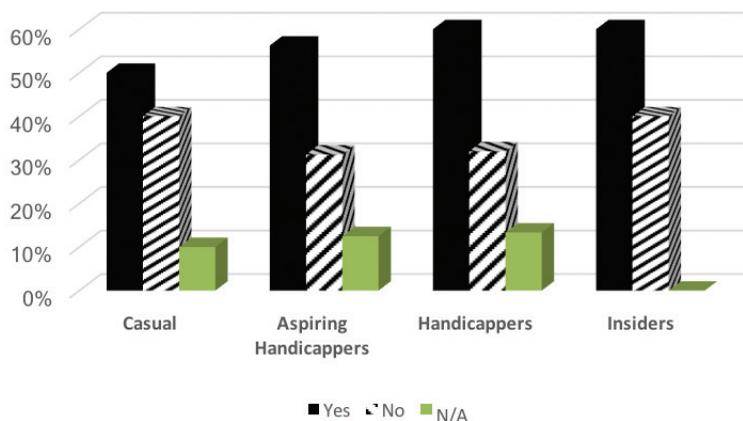
of the day in an effort to recoup losses. Therefore, I was surprised to see 77% of bettors say they would remain consistent, and only 11% say they would increase their wagers. I believe respondents were trying to answer the question logically, instead of admitting they would let their emotions get the best of them. Increasing the amount you bet is not a rational response, yet in the heat of the moment feeling your pockets get lighter and lighter your judgment is clouded by your desire to go home a winner.

The last question on the survey dealt directly with the overall theme of my study. The market is efficient according to academic research, but is it efficient according to the public? I found that 57% of the respondents believe it is possible to consistently ‘beat the horse racing market,’ thus making this market inefficient. In contrast, 34% of the participants disagree, indicating it is efficient. The remaining 9% provided no answer, perhaps because they did not understand the question or did not feel strongly either way. Figure 7 breaks these answers down by betting class in order to properly gauge what influences people to gamble on thoroughbred horses.

The casual bettors were fairly even in their beliefs with 50% answering yes, 40% answering no, and 10% having no response. I believe these statistics represent the causal betting population well, for there are various motivations driving these wagers; therefore some believe the market is beatable while others do not. I consider myself a casual bettor, and as such, I do not believe it is possible to consistently beat the market using my method of wagering. Yet, I go to Saratoga Racetrack every year with my family mainly for the experience, but also because I believe in the small chance I will hit a longshot and go home a winner. I also find that I am attracted to the thrill of potentially holding a winning ticket. I bet because then I have a stake in the race—a horse to root for. I believe the casual bettors who answered “no” share similar reasons for betting. At the same time, I believe those who do enough research and have access to privileged information are able to consistently beat the horse racing market. They will not win every wager they make, but they can earn above average profits through their informational advantage.

FIGURE 7

Is it Possible to Beat the Markets? (Is the Market Inefficient?)



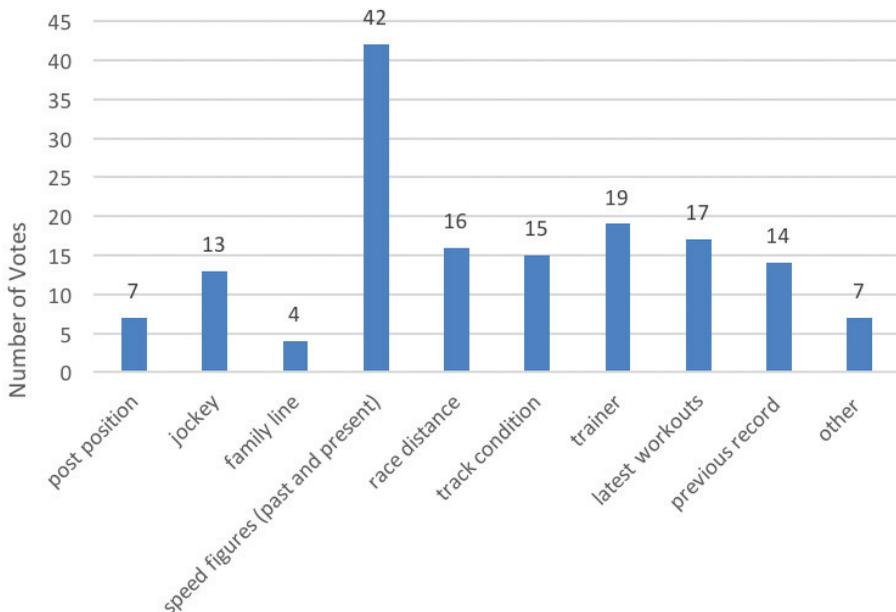
Of the aspiring handicappers, 56% of them believe the market is inefficient, 31% believe it is efficient, and 13% did not provide an answer. Logically, it makes sense for these bettors to believe the market is inefficient, otherwise, their beliefs contradict one another. In other words, by indicating they believe the market is efficient these bettors are saying they want to become more informed by studying horse racing data, but doing so will still not prove successful. Why put effort into something you do not believe you can win anyway? For some the few victories outweigh the several losses. One could ask baseball players a similar question. Why do they step into the batter's box day in and day out knowing they will fail seven out of ten times? For a baseball player, the adrenaline-pumping thrill he experiences hitting the ball on the sweet spot of the bat is worth all the failure. Betting at the track is similar for many aspiring handicappers. The excitement one feels when he sees his horse crossing the finish line first, overshadows the seven races he lost before that. Aspiring handicappers believe they can experience such a thrill more often if they become more informed bettors.

Of the handicappers, 60% of them believe it is possible to beat the market, 32% of them do not, and 8% provided no answer. Similar to the aspiring handicappers, the surprising statistic here is the number of professional handicappers who believe that which they devote countless hours to is not actually possible. Many of these bettors pour thousands of dollars into horse racing, but for what? Some might be addicted to the feeling of hitting a big winner. Others may believe it is not possible to beat the market, but find it a challenge trying to prove otherwise. Exploring the reasons behind their answers offers some interesting insight that parallels points I have discussed throughout my review of the literature. One respondent claimed it is possible to consistently beat the market, but "it is getting harder to do so" (survey). In essence, this bettor feels that handicappers are losing their competitive advantage, perhaps because their betting strategies are outdated. Their methods are being discovered and used by more people, reducing their effectiveness. This is the same point Beyer portrays through *Beyer on Speed*. Additionally, other bettors claimed the market is efficient because "the track take is too high" (survey). It becomes very difficult to earn a worthwhile profit when almost 20 cents of every dollar one wins is going to the racetrack. As explained in the literature review, this is why the favorite-longshot bias is considered an unprofitable anomaly, as opposed to proof of the market's inefficiency. Why these handicappers continue to bet so frequently and with large sums of money remains unknown. We can only surmise and recognize that people are driven by a series of conflicting emotions, which affect their betting behavior.

At the same time, the 60% of handicappers who believe it is possible to consistently beat the market offer meaningful insight through their comments. They explain it is possible to beat the market, but one must be disciplined and patient. In order to succeed as a bettor, you must devise an informative betting strategy and pursue it. Discipline means refraining from betting on the 50 to 1 shot simply because the thought of winning on that horse excites you. Patience means not betting on every race, but only those you believe offer a competitive advantage. One of the major differences between

causal bettors and handicappers is the former typically bet on every race at the track, whereas the latter will not bet unless they believe there is an inefficiency to be exploited. As advised by one of the handicappers surveyed, “You have to handicap and only bet the race or races you have a strong opinion on” (survey). In addition, serious bettors only focus on those bets, which will be profitable for them. They will not bet on a horse that everyone else also believes will win because it is not lucrative. For instance, one of the insiders claims “dumb money is only around on big days” (survey). The term ‘dumb money’ refers to the casual bettors’ wagers. Casual bettors do not bet with any sound information or strategy, which causes more inefficiencies in the market. Hence, on weekends and in the summertime there will be more profitable betting opportunities for the handicappers due to the high attendance among causal bettors. Accordingly, it might be surprising to some that 15 of the 30 serious bettors wager on longshots most often. The key distinction being that handicappers only bet on ‘longshots’ with potential that goes unnoticed by the public. According to their handicapping strategies or inside knowledge such horses are not actually longshots, yet their publically determined odds dictate so. Most longshots bet by handicappers are mispriced investments, as Beyer would say. This betting class understands they are not going to win every time they wager; however, a majority of them believe a handicapper can earn above average profits as long as they place sound bets, backed by valuable information and react appropriately to the wisdom of the crowd.

**FIGURE 8**  
Bettor’s Most Important Three Data Elements



Since the use of historical data is so abundant in horse racing, I found it would be beneficial to know what type of past information handicappers consider the most advantageous. In my survey, I asked participants to designate their top three most important elements they consider when picking a winning horse. The most common answer was speed figures with 42 selections, as shown in Figure 8. Another popular, yet unexpected choice was trainer with 19 selections. However, with more thought this seems logical due to trainers' unique workout styles. A horse must have natural speed and skill, but the way he is trained can make him a champion. Many of these elements such as, jockey, race distance, track condition, latest workouts, and previous record, received equal interest with around 15 voters each. Family line and post position do not seem to foster as much awareness from bettors. I was surprised to see low regard for post position because where a horse is placed in the starting gate can greatly impact a race. Often times a horse who is supposed to win does not because it did not get an advantageous start out of the gate. However, taken as a whole this data demonstrates the copious amounts of information the horse racing market supplies. Furthermore, it shows that bettors place different weights on different elements. As a bettor, it is up to you to determine what elements of information are the most important.

When asked to reveal the underlying basis of their bets, 29% of bettors said Beyer speed figures, 9% said they mirror the picks of professional handicappers, and 2% admit they select the horse with the coolest name. 61% selected 'other' as their answer choice and then explained the main strategy they use to pick a winning horse. A number of respondents indicate they calculate their own speed figures using all the information available to them. This is fascinating and demonstrates the dedication of many of these serious bettors. Another surveyor said he had a private database, which assists him in calculating a horse's actual probability of victory. Some bettors indicate they use a variety of past performance and pace figures. In essence, the information is endless. There are countless methods to picking a winning horse. It comes down to personal preference and developing a strategy that is protected from use by the general public.

Despite the seemingly infinite supply of horse racing data, even a horse you are more than confident about may not be victorious at the end of the day, which makes horse betting so captivating. One may be able to demonstrate consistent profits, but it is impossible for them to win every time. For this reason, I asked respondents to tell me about a moment when all of their data pointed to a certain horse winning and it did not. Their responses were eye-opening. Many of them explained the horse's inadequate performance was due to poor position, a bad trip, being overrun, unusual track conditions, or the jockey not riding the horse well. Others contend "a horse is an animal, not a machine;" therefore mistakes will be made (survey). This sport, similar to others, is not black and white. At the end of the day, no matter how much information you use to forecast a winner you still need luck on your side.

## V. Conclusion

As demonstrated, the stock market and the horse racing market offer some illuminating points of comparison. Although most of the literature proclaims the horse racing market efficient, many of the bettors who travel to the track would disagree. Interestingly enough, these conflicting views help the market prevail and compel the crowd to place their bets. Despite behavioral impact, the subjective odds for each horse which are determined by the wisdom of the crowd, come close to mimicking each horse's empirical odds. This indicates the market is efficient at least in the weak form. An anomaly known as the favorite-longshot bias, in which bettors tend to underbet favorites and overbet longshots, poses a threat to market efficiency. However, this anomaly is often not reliable enough to overcome track take. In essence, this short-lived anomaly is inconsequential to prove inefficiency in the horse racing market similar to several of the anomalies in the stock market such as, the small firm effect and the January effect. Most of the research indicates the market is also efficient in the semi-strong form, but several handicappers demonstrate otherwise through their lucrative profits. In the semi-strong form, handicappers display adequate betting strategies that generate above-average returns, but these tactics begin to lose their competitive advantage over time as public awareness grows. Literature on the strong form is minimal because evaluating private information would expose it and diminish its benefit. However, it is reasonable to believe insiders such as jockeys, trainers, and owners, who work directly with the horse, have access to privileged information that will assist them in the betting arena. In the end, the horse racing market, similar to the stock market, is efficient in the long term for casual bettors and investors. However, for those handicappers who carefully incorporate an abundance of market information into their bets and for those insiders who hold access to privileged information, the market demonstrates significant inefficiencies that they can exploit.

As a casual bettor, there are many lessons to be learned through the analysis of market efficiency. To begin, because the horse racing market is efficient in the long-term one should not bet on every race. Instead, casual bettors should try to incorporate as much available information into their picks as possible in order to make an educated wager, and only place a bet when they see a horse that is mispriced. Our behavior plays a significant role in gambling, whether we like it or not. We bet on that longshot because of the excitement we get thinking of cashing in a money-spinning ticket. We overlook that favorite because we did not come to the track to make \$5, but to make \$50 or \$500. Through these behaviors horses become mispriced, similar to the way securities in the stock market become mispriced, for "stocks are usually more than just the abstract 'bundle of returns' of our economic model" (Merton Miller as quoted in Thaler and Ziemba, 1988). Behind each bettor's wager is a story or sentiment, unknown to the public. It might be a story of the first day at the track, the last bet of the summer racing season, the need to impress wealthy business colleagues, or even a wager backed by a lot of liquid courage. Whatever the story may be, the betting be-

haviors that follow will impact the market causing inefficiencies. A casual bettor with the desire to be profitable needs to be rational, separate herself from these sentiments, and focus on the horse racing data. However, an interesting paradox exists in the horse racing market, which then comes into play. As more and more bettors probe for information to make educated bets, the chance for profitability declines. If more people are informed, then horses will be adequately priced, reducing exploitation opportunities. In essence, we are running our own race—around and around in circles we wager.

## References

- Asch, P., Malkiel, B. G., & Quandt, R. E. (1982).** “Racetrack Betting and Informed Behavior.” *Journal of Financial Economics* 10(2): 187-194.
- Beyer, Andrew. (1993).** *Beyer on Speed*. Boston: Houghton Mifflin.  
Daily Racing Form. Retrieved from <http://www.drf.com/>
- Daily Racing Form. (May 16, 2016).** Volume 120. No. 136. (NY)
- Degutis, A., & Novickyte L. (2014).** “The Efficient Market Hypothesis: A Critical Review of the Literature and Methodology.” *Ekonomika*, 93(2): 7-20.
- Douglas, N. (1995).** “In the light of current evidence, critically examine the efficient market hypothesis.” Retrieved from <http://www.nedprod.com/studystuff/MN3101%20Essay.pdf>
- Durham, G. R., Hertz, M. G., & Martin, J. S. (2005).** “The Market Impact of Trends and Sequences in Performance New Evidence.” *The Journal of Finance*, 51(5): 2551-2569.
- Fama, E. (1965).** “Random Walks in Stock Market Prices.” *Graduate School of Business, University of Chicago, (selected paper No. 16)*.
- Fama E. (1970).** “Efficient Capital Markets: A Review of Theory and Empirical Work.” *Journal of Finance*, 25(2): 383-417.
- Fluck, Z., Malkiel, B. G., & Quandt, R. E. (1997).** “The Predictability of Stock Returns: A Cross-Sectional Simulation.” *The Review of Economics and Statistics*, 79(2): 176-183.
- Gramm, M., & McKinney, C. N. (2009).** “The Effect of Late Money on Betting Market Efficiency.” *Applied Economics Letters*, 16: 369-372.
- Hausch, D. B., Ziemba, W. T., & Rubinstein, M. (1981).** “Efficiency of the Market for Racetrack Betting.” *Management Science*, 27(12): 1435-1452.
- Kissell, R. (2014).** “Belmont Stakes, Stanley Cup Final Score Ratings Milestones for NBC Sports on Saturday.” *Variety*.

- Losey, R. L., & Talbott, J. C. (1980).** “Back on Track with the Efficient Markets Hypothesis” *Journal of Finance* 35(4): 1039-1043.
- Malkiel, B. G. (2003).** “The Efficient Market Hypothesis and Its Critics.” *Journal of Economic Perspectives*, 17(1): 59-82.
- Mauboussin, M. J., & Callahan, D. (2015).** “Triple Crown Teachings: Three Takeaways from the Quest for the Elusive Feat.” *Credit Suisse*.
- McGrath, M. (2015).** “Belmont Stakes 2015 by the Numbers: Quantifying a Triple Crown Attempt.” *Forbes/Investing*.
- Mishkin, F. S. (2012).** Appendix to Chapter 7: Evidence on the Efficient Market Hypothesis. In *The Economics of Money, Banking and Financial Markets*. Pearson.
- Ottaviani, M. & Sorensen, P. N. (2009).** “Surprised by the Pari-mutuel Odds?” *American Economic Review*, 99(5): 2129-2134.
- Rosenbloom, E. S. (2003).** “A Better Probability Model for the Racetrack Using Beyer Speed Numbers.” *Omega: The International Journal of Management Science*, 31: 339-348.
- Singh, L. K. (2010).** “Wither Efficient Markets? Efficient Market Theory and Behavioral Finance.” *The Financial Professionals’ Post*.
- Snowberg, E. & Wolfers, J. (2010).** “Explaining the Favorite-Longshot Bias: Is it Risk-love or Misperceptions?” *Journal of Political Economy*, 118(4): 723-746.
- Snyder, W. W. (1978).** “Horse Racing: Testing the Efficient Markets Model.” *The Journal of Finance* 33(4): 1109-1118.
- Thaler, R. H., & Ziemba, W. T. (1988).** “Anomalies: Parimutuel Betting Markets, Racetracks and Lotteries” *Journal of Economic Perspectives* 2(2): 161-174

**Appendix***HORSE RACING SENTIMENTS SURVEY*

1. What kind of horse race bettor do you consider yourself to be?
  - Casual bettor- I want to win money, but I don't look up betting strategies
  - Aspiring handicapper- There is data-driven meaning behind my bets. I try to implement a strategy, but I would not consider myself a regular handicapper
  - Handicapper- I devote hours to studying the DRF and various betting strategies
  - Insider- I am a trainer, jockey, owner, etc. and have access to restricted information
  
2. How many times per year do you usually bet on horse races?
  - once or twice a year
  - once a week
  - once a month
  - 3 times a month
  - Other (please specify)
  
3. Are you a tournament bettor? If so, which tournaments do you enter?
  
4. With how many minutes to post (MTP) do you usually place your bet?
  - 20-24 minutes
  - 15-19 minutes
  - 10-14 minutes
  - 5-9 minutes
  - Less than 5 minutes
  
5. About how much do you typically bet on horse races each year?
  - Less than \$100
  - \$101 to \$300
  - \$301 to \$500
  - \$501 to \$1000
  - \$1001 to \$5000
  - \$5001 to \$20,000
  - More than \$20,000

6. Rank each bet type in your order of preference
  - Win
  - Place
  - Show
  - Across the board (win, place, show)
  - Exacta
  - Trifecta
  - Daily Double
  - Pick 3
  - Pick 6
  
7. What type of horses do you find yourself betting on most often?
  - Favorite
  - Longshots
  - Neither
  
8. Please tell us about your largest victory in thoroughbred horse betting.
  
9. How do you choose the horses you bet on?
  - I choose the horse with the coolest name
  - I choose a horse based on the number it will be wearing
  - I used Beyer speed figures
  - I pick my horse based on the professional handicappers' picks noted inside the DRF
  - Random (i.e.: I put the numbers in a hat and pick)
  - Other (please specify)
  
10. If possible, please describe a time when handicapping methods proved beneficial.
  
11. If possible, please describe a time when all the data pointed to a particular horse winning and that horse did not win. Why do you think this horse was not victorious?
  
12. It is the eighth (and final) race of the day. You have bet on all the previous races, and you have not yet won. You have lost a fair amount of money on your previous bets. What do you do in the eighth race?
  - Don't place a bet at all on the last race
  - Wager less than you have been all day
  - Wager an amount consistent with the amount you have been betting all day
  - Increase the amount of this last bet or bet on a longshot in the hopes of winning big

13. What data elements do you find are the most important when trying to select a winning horse (please do not select more than 3):

- Speed figures
- Track condition
- Post position
- Jockey
- Trainer
- Family line
- Previous record (first places wins, second, and third)
- Latest workouts
- Race distance
- Other (please specify)

14. Would you consider the horse race betting market to be an example of an efficient market? In other words, do you believe it is not possible to consistently earn above average profits? Please explain.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the contributions of our reviewers in supporting *Fusio* and the authors in ensuring the highest standards of academic quality. The authors have benefitted tremendously from their time and insightful comments, so we thank you.

In particular, we are indebted to the following reviewers for Volume I of *Fusio*:

Elena Abascal	Helen Meldrum
Matthew C. Ahlmeyer	Javier Monterrosa Hernandez
Mystica Alexander	Jim Musumeci
Liam Bartlett	Jeffrey Pierce
Chris Beneke	Sarah Pourshadi
Natalie Cotton-Nessler	Jeff Proudfoot
Marc Custodio	Michael Quinn
Dhaval Dave	John Rogers
Gary David	Christian Rubio
Virginia Duffy	Sydney Sachs
Matt Dwyer	Eshita Shah
Hans Eijmberts	Jeremy Siborg
Otgontsetseg Erhemjamts	Kristin Sorensen
Juliet Gainsborough	Euthemia Stavoulaki
Jeff Gulati	Angela Su
Dave Gulley	David Szymanski
Charlie Hadlock	Marie Tuchy
Danielle Hartigan	Dulguun Tuvshintugs
Bruce Herzberg	Katie Vadakin
Emily Hunt	David Yates
Aaron L. Jackson	Laura Young-Jackson
Fred Ledley	Mateusz Zeglen
Michael Liotti	Josh Zheng

## ARTICLES

Bridging the Gap: Securing the Internet of Things

**Matthew C. Ahlmeyer**

The Implications of Religious Freedom Restoration Laws and the Evolution of Free Exercise Protection in the United States

**Amanda Pine**

Managing Stigma: Women Drug Users and Recovery Services

**Nayeong Lee and Miriam Boeri**

Yay or Neigh? Is the Horse Racing Market Efficient?

**Jillian Raia**

Acknowledgements

**Bentley University Honors Program**

Waltham, MA, USA

fusio@bentley.edu



**BENTLEY**  
UNIVERSITY