2015

Incomplete Reporting: Addressing the Problem of Outcome-Reporting Bias in Educational Research

Brian Patrick Trainor

Loyola University Chicago

Recommended Citation


http://ecommons.luc.edu/luc_diss/1656

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.

Copyright © 2015 Brian Patrick Trainor
LOYOLA UNIVERSITY CHICAGO

INCOMPLETE REPORTING:
ADDRESSING THE PROBLEM OF OUTCOME-REPORTING BIAS IN
EDUCATIONAL RESEARCH

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE GRADUATE SCHOOL
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

PROGRAM IN RESEARCH METHODOLOGY

BY
BRIAN P. TRAINOR
CHICAGO, IL
AUGUST 2015
Copyright by Brian P. Trainor, 2015
All rights reserved.
ACKNOWLEDGMENTS

This dissertation could not have been possible without first the love and support of my parents, William and Janice. To William, who demonstrated to me not only the value of hard work, but also the capacity a person can have for continued, selfless dedication to one’s family; and to Janice, whose gentle heart and unwavering love have served as an guiding light to me: I thank you.

My tremendous thanks also go out to Loyola University Chicago for providing me with the assistantship that has allowed me to continue my scholarship in the field of education and research methodology. The use of Loyola’s facilities and resources has been an integral part of this process.

The faculty and staff at Loyola have also served as amazing resources throughout my graduate studies. My deepest thanks go out specifically to Dr. Terri Pigott, who gave me the honor of serving as my advisor. There is not a price one could put on the amount of advice, time, guidance, and insight Dr. Pigott has provided. It is my hope that long after my time at Loyola has ended; Dr. Pigott will continue to serve as my mentor, colleague, and dear friend.

I would also like to thank the fellow Loyola students who had been going through this process with me; specifically William Adams, Brendan Martin, and John Segvich. These individuals have not only offered their unique insights during our many classes together, but they also allowed me to run ideas by them and served as great friends.
Many thanks are also in order for my brother Michael. Not only does he challenge me to continually question the things around me in pursuit of a deeper understanding of the world, he also always provided me with the necessary distractions when the daunting task of writing a dissertation loomed nearer and nearer. Most importantly, I would like to thank him for his humor, and for instilling in me a lifelong obsession with comic books.

Finally, and most importantly, I would like to thank my wife, Lauren O’Connor. Lauren’s love has allowed me to become myself. She has shown me who I was, who I am, and who I was always meant to be. She encourages me to pursue my passions and to live up to my potential. She is the most intelligent and toughest person I know. Lauren continually serves as my inspiration and I could not have completed this process without her. To my wife Lauren: You are the best thing that ever happened to me.
For Lauren
There are three kinds of lies: lies, damned lies, and statistics.

—Mark Twain
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER ONE: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Research Questions</td>
<td>4</td>
</tr>
<tr>
<td>Significance of the Study</td>
<td>6</td>
</tr>
<tr>
<td>CHAPTER TWO: LITERATURE REVIEW</td>
<td>9</td>
</tr>
<tr>
<td>Outcome-Reporting Bias</td>
<td>10</td>
</tr>
<tr>
<td>Outcome-Reporting Bias in Educational Research</td>
<td>17</td>
</tr>
<tr>
<td>Factors Influencing Outcome-Reporting Bias</td>
<td>20</td>
</tr>
<tr>
<td>Publication Bias</td>
<td>21</td>
</tr>
<tr>
<td>Importance of Meta-Analysis and Systematic Review</td>
<td>26</td>
</tr>
<tr>
<td>Potential Threats to Validity</td>
<td>28</td>
</tr>
<tr>
<td>CHAPTER THREE: METHODS</td>
<td>30</td>
</tr>
<tr>
<td>Established Methods of Comparison</td>
<td>31</td>
</tr>
<tr>
<td>Selecting Research Institutions in the Original Study</td>
<td>32</td>
</tr>
<tr>
<td>Search Strategies in the Original Study</td>
<td>32</td>
</tr>
<tr>
<td>Identifying Subsequent Articles in the Original Study</td>
<td>33</td>
</tr>
<tr>
<td>Reliability Checks and Replicated Searcher</td>
<td>34</td>
</tr>
<tr>
<td>Coding and Data Analysis</td>
<td>35</td>
</tr>
<tr>
<td>CHAPTER FOUR: RESULTS</td>
<td>44</td>
</tr>
<tr>
<td>Do Differences Exist Between the Dissertations and the Articles?</td>
<td>45</td>
</tr>
<tr>
<td>Outcome-Reporting Bias and Statistical Significance</td>
<td>47</td>
</tr>
<tr>
<td>The Effect of “Positiveness”</td>
<td>48</td>
</tr>
<tr>
<td>Results Consistent with the Authors’ Hypotheses</td>
<td>49</td>
</tr>
<tr>
<td>Academic Affiliations and Outcome-Reporting Bias</td>
<td>51</td>
</tr>
<tr>
<td>Other Potential Factors Identified</td>
<td>53</td>
</tr>
<tr>
<td>CHAPTER FIVE: DISCUSSION</td>
<td>56</td>
</tr>
<tr>
<td>Differences that Exist Between Dissertation/Article Pairs</td>
<td>56</td>
</tr>
<tr>
<td>Outcome-Reporting Bias and Statistical Significance</td>
<td>59</td>
</tr>
<tr>
<td>The Effect of “Positiveness”</td>
<td>61</td>
</tr>
<tr>
<td>Results Consistent with the Authors’ Hypotheses</td>
<td>62</td>
</tr>
<tr>
<td>Academic Affiliations and Outcome-Reporting Bias</td>
<td>63</td>
</tr>
<tr>
<td>Other Potential Factors, Final Thoughts, and Suggestions for Future Research</td>
<td>64</td>
</tr>
<tr>
<td>APPENDIX A: CODING INSTRUMENT</td>
<td>67</td>
</tr>
</tbody>
</table>

vii
LIST OF TABLES

Table 1. Coding Information for Study and Outcome Identification 36
Table 2. Coding Information Regarding Statistical Significance and Reporting 36
Table 3. Coding Information for Harmful Effects 37
Table 4. Coding Information for Consistency 38
Table 5. Data Coding Information for Academic Affiliations 39
Table 6. Data Collected From Each Dissertation/Article Pair For Use in Question 2a. 41
Table 7. Data Collected From Each Dissertation/Article Pair For Use in Question 2b. 41
Table 8. Data Collected From Each Dissertation/Article Pair For Use in Question 2c. 42
Table 9a. Data Collected From Each Dissertation/Article Pair For Use in Question 2d. 42
Table 9b. Data Collected From Each Dissertation/Article Pair For Use in Question 2d. 42
Table 10. Descriptive Statistics for Dissertation/Article Pairs 45
Table 11. Reporting Statistically Significant Outcomes 46
Table 12. Results Consistent with the Authors’ Hypotheses 50
Table 13. Academic Affiliations Subgroup Analysis 52
Table 14. Grade Level Subgroup Analysis 54
Table 15. Race Subgroup Analysis 54
ABSTRACT

Outcome-reporting bias is a problem that pervades many research disciplines including education. Outcome-reporting bias involves any time the outcomes presented in a published journal article do not accurately reflect all of the outcomes that were collected throughout the course of a study. In other words, when outcome-reporting bias is present, the information which is disseminated to the academic research community is incomplete and can lead to serious problems over time. Some of these problems include school districts implementing interventions based on incomplete or inaccurate data, as well as the problem that naturally arises for those implementing meta-analysis or systematic review research strategies. Many times these important research methodologies rely on compiling and synthesizing all outcomes that have been collected across a number of different studies. If not all outcomes were made available to the public, the conclusions derived from these important types of studies would be invalid and/or based on potentially biased data. In order to ensure that educators and educational researchers alike are implementing truly effective interventions and that the conclusions derived from meta-analyses and systematic reviews are valid, exploration into factors associated with outcome-reporting becomes necessary. The purpose of this dissertation project was to explore the nature of outcome-reporting bias by identifying potential factors associated with it. This task was achieved by first exploring instances in which outcome-reporting bias was documented in other academic research fields and
determining what factors were associated with it in those particular cases. The next step became thoroughly reviewing educational research articles and determining which of those factors appeared to be associated with outcome-reporting bias in educational research. The hope is that by identifying these factors, researchers and journal publishers can work to eliminate them from practice, decreasing the prevalence of outcome-reporting bias in educational research. This study found that not only is outcome-reporting bias present in educational literature, but also that outcomes failing to show statistical significance were 30% more likely to get suppressed than statistically significant outcomes, and outcomes that were not consistent with the publishing author’s original hypothesis were 41% more likely to get suppressed than those that were. In addition, this study also found that researchers holding both faculty and non-faculty positions appeared to exhibit outcome-reporting bias with greater frequency when it comes to statistical significance. Non-statistically significant outcomes were 26% more likely to get suppressed than statistically significant outcomes among individuals holding faculty positions and 50% more likely among non-faculty researchers. These results show that authors have a tendency to withhold outcomes that were not statistically significant as well as those that were not consistent with their initial hypotheses. In addition, there must be some type of pressure present that leads to researchers withholding non-statistically significant outcomes from publication. Further, this study also found that data collected regarding certain populations are more subject to outcome-reporting bias when it comes to statistical significance. When samples are predominantly white, non-statistically significant outcomes are 24% more likely to be suppressed when compared to 73% among predominantly non-white samples. Also, non-statistically significant
outcomes are 25% more likely to be withheld among high school samples and 32% more likely to be withheld among non-high school samples.
CHAPTER ONE

INTRODUCTION

Outcome-reporting bias exists when researchers submit papers for publication but fail to include information about all of the outcomes that were measured during the study. In order to better understand what outcome-reporting bias looks like, picture an individual who is performing research on a newly developed academic intervention. During the data collection phase the researcher collects information on both math and reading outcomes associated with the intervention. If the researcher were to write up the results of this study and submit it for publication but only include the math outcomes associated with the intervention while selectively excluding the reading outcomes, outcome-reporting bias would be present. This scenario is only one example of a way in which outcome-reporting bias can exist. While it seems extreme, one could argue that if a researcher were to exclude any single outcome from the final publication then a form of outcome-reporting bias would exist. Although this practice seems highly unethical, a thorough review of the literature will show that not only does outcome-reporting bias exist, but it also pervades many research disciplines.

The selective inclusion and/or exclusion of intervention outcome data can have a tremendous impact on the academic community. One issue that can arise if authors choose to selectively report/withhold outcome data is that not all of the information about a particular intervention’s effectiveness would ultimately be disseminated to the public.
This problem can lead to individuals, companies, or organizations implementing interventions or making policy decisions without first having all of the necessary information. It is impossible to make an informed decision regarding whether or not an intervention or policy should be implemented if certain outcomes were deliberately withheld from publication. This issue also creates a particular source of concern for any future researchers who wish to perform a meta-analysis or systematic review on the effectiveness of the intervention in question because they would unknowingly be working with an incomplete data set.

Performing a meta-analysis or systematic review on an intervention would involve compiling every outcome that has ever been collected on it and combining the results in order to provide a synthesized conclusion with greater statistical power. Further, meta-analyses and systematic reviews are used to make more informed claims regarding an intervention’s overall effectiveness. If not all of the outcomes that were collected on the intervention were made available due to the problem of outcome-reporting bias, any meta-analysis or systematic review performed on the intervention would be essentially useless. In fact, the subsequent meta-analysis or systematic review could potentially compound the existing bias on a larger scale.

Given that the availability of and access to information, specifically outcome data, is important for meta-analysis research, it is vital that nothing be withheld from the research community and that full transparency becomes an expectation. Despite the essential nature that is accessibility/transparency, it has been well documented that researchers continue to withhold certain information from their publications due to a number of biases, including, but not limited to, outcome-reporting bias. The different
types of biases that affect this perceived lack of transparency are detailed in a thorough review of the literature later in this dissertation.

What makes the problem of outcome-reporting bias even more complicated is that it is not always possible to discover whether or not it was even present within the literature. This potential lack of access to outcome data can pose a major threat to the validity of the results derived from meta-analyses and systematic reviews, making it all the more important for the academic community to establish ways to better identify whether or not outcome-reporting bias is present as well as what factors are associated with it. Once this important task is accomplished the scientific community can begin taking steps to potentially eliminate these factors from practice.

Just as methodologists have devised ways to identify other types of bias within scholarly research, ways to deduce whether or not outcome-reporting bias is present are also beginning to emerge. In fact, outcome-reporting bias has already been identified in a number of different disciplines. That being said, researchers in many of the social sciences have only begun taking the steps necessary to discover the degree to which outcome-reporting bias has pervaded their respective fields (Torgerson, 2006), and it was not until recently that work has been done to explore its prevalence in educational research.

Pigott, Valentine, Polanin, Williams, and Canada (2013) were essential in exploring the existence of outcome-reporting bias in educational research. The authors found that outcome-reporting bias was indeed present within the educational literature and that researchers were more likely to exclude non-statistically significant outcomes from their final publications than those that showed statistical significance. While Pigott
et al. (2013) documented the existence of outcome-reporting bias in educational research and gave the scientific community an initial glimpse into a potential reason why researchers may choose to exclude certain outcomes from their final publications, one may ask if the lack of statistical significance is the only factor associated with outcome-reporting bias or if there are others that need to be considered.

Other factors that may be associated with a researcher’s tendency to exhibit outcome-reporting bias can be complex and potentially rooted in other forms of bias like publication bias. These factors, when coupled with other forms of bias, can produce a vicious cycle that ultimately exacerbates the overall effect bias has on academic research. Further, and perhaps more disastrous, this cyclical nature of bias can lead to the research community implementing interventions based upon literature that withheld important outcome data from the scientific community. One way to help circumvent this cycle is to identify the factors that are associated with outcome-reporting bias in educational research and suggest ways to eliminate them from scholarly research practices. It is the purpose of this dissertation to identify such factors in order to get a better understanding of outcome-reporting bias and its prevalence in educational research so as to encourage all researchers and publishing editors alike to be forthright with their work and to publish all outcomes that were collected on an intervention.

Research Questions

The work presented here will expand upon the work done by Pigott et al. (2013) by identifying some factors that may be associated with outcome-reporting bias and suggest ways to eliminate them from practice in order to better disseminate valid claims to the educational research community. Exploring the nature of outcome-reporting bias as
well as identifying potential factors associated with it can be achieved through an analysis of outcomes presented in an individual’s dissertation to those in subsequent journal articles. The main focus of this dissertation can be best expressed through the understanding of, and subsequent answers to, the research questions presented below:

1. Do differences exist between the outcomes reported in an individual’s doctoral dissertation and those reported in subsequent publications utilizing the same data set?

2. If outcome-reporting bias exists in the identified sample of dissertations and published articles, what factors are associated with its presence?

   a. Are statistically significant outcomes more likely to get reported than those that are not statistically significant?

   b. Are statistically significant positive outcomes more likely to get reported than those that showed statistically significant negative outcomes or harmful effects associated with a treatment?

   c. Are outcomes consistent with the author’s initial hypothesis more likely to get reported than those that are not?

   d. Do authors having pursued academic careers in which publication is necessary have higher instances of outcome-reporting bias than those that did not?

   e. Are there any factors associated with sample demographics that lead to higher instances of outcome-reporting bias?

In order to best answer the aforementioned questions a systematic review of educational literature was performed. The specific methods used in the data collection process will be discussed later in this dissertation and were based upon those designed by Pigott et al. (2013) in which the outcomes found within doctoral dissertations were
compared to those found within subsequent journal publications using the same data sets. This time, however, the search was repeated and additional information was retrieved and coded from each study so as to ascertain not only if outcome-reporting bias exists, but also what factors might be associated with its presence. It should be noted that much of the search process used to identify the original dissertation-publication pairs in the original study by Pigott et al. (2013) was replicated for the purposes of this study.

**Significance of the Study**

The use of meta-analysis and systematic review among researchers and policymakers to facilitate exploration and guide decision making continues to increase dramatically (Field, 2003). This increase in the utilization of meta-analysis and systematic review methods can be found across many different fields of research including, but not limited to, medicine (Higgins & Green, 2011), psychology (Cooper, 2010), and education (Pigott, 2012). Given this shift in the field of scientific inquiry, it is important for the academic community to better understand meta-analysis and systematic review as research methodologies as well as the factors that may influence or affect the validity of the data provided within these studies.

Not only should researchers be educated on the intricacies of employing these research designs, it is also vital that these individuals understand what potential threats to validity exist within the methodologies themselves. Outcome-reporting bias is one such threat to the validity of primary research and syntheses that could dramatically affect not only the type of information that is disseminated to the research community but also how well the fields of meta-analysis and systematic review are perceived by researchers as reputable and relevant methodologies.
This study will primarily build upon the work done by Pigott et al. (2013) by replicating their study in order to ensure nothing was missed in gauging whether or not outcome-reporting bias has pervaded educational literature. A second aim of this study is to establish what factors may be associated with the presence of outcome-reporting bias among researchers. This study is significant to many academic fields including education, research methodology, and meta-analysis and systematic review. Educators need to be made aware of the existence of outcome-reporting bias before they begin implementing policy decisions or academic interventions based on research that failed to include all of the collected outcome measures. Research methodologists and researchers alike need to better understand the factors that lead to outcome-reporting bias so they can improve their practices, avoid these factors, and ultimately disseminate valid claims to the scientific community as well as meta-analysts and systematic reviewers. Those in the fields of meta-analysis and systematic review need to be made aware of ways to identify the presence of outcome-reporting bias because it serves as a potential threat to the validity of their work.

The work contained within this study will serve as a call to action for individual researchers, as well as publishing editors and research institutions alike, that outcome-reporting bias, as well as the factors that may be associated with it, can greatly affect the types of information that is presented to the scientific community and the quality of research that exists. These factors must be identified so they can be minimized, if not eliminated. If the educational research community can better understand whether or not outcome-reporting bias has pervaded the field as well as identify what factors might be associated with it, they can work to establish practices aimed at ensuring all outcomes
collected on interventions can be made available and fully informed decisions regarding interventions can be made. Increasing the level of accessibility to and reporting of all outcome data also ensures individual researchers utilizing important methodologies like meta-analysis and systematic review can be confident that the findings derived from their work are valid. It is only when the findings are valid that the educational community can claim an intervention is truly effective and appropriate for implementation.
CHAPTER TWO

LITERATURE REVIEW

The primary aim in this dissertation is to present the academic community with a better understanding of outcome-reporting bias in the field of education—its existence in educational research, the issues surrounding it, the subsequent impact it has on the fields of systematic review and meta-analysis, and what potential factors are associated with it. The work presented here will serve to raise awareness among primary researchers, journal editors, and those performing systematic reviews and meta-analyses in the field of education of the problem that is outcome-reporting bias. Addressing this problem will allow for researchers to decrease its prevalence as well as lead to better research practices and increased validity among the findings derived from systematic reviews and meta-analyses.

The contents of this chapter will, through a thorough review of the literature, detail what outcome-reporting bias is and identify what potential factors may be associated with such a bias. Pinpointing these factors will be achieved by exploring instances in which outcome-reporting bias was documented in other fields of academic research and then identifying what factors appeared to be associated with outcome-reporting bias there. This chapter will also show the ways in which outcome-reporting bias can influence the validity of subsequent systematic reviews and meta-analyses. Finally, this chapter will document exactly how other well-documented forms of bias can exacerbate the prevalence of outcome-reporting bias within educational research.
Outcome-Reporting Bias

Outcome-reporting bias occurs when published primary studies fail to include information covering all of the primary or secondary outcomes that were originally measured/collection (Pigott, et al., 2013). Studies that exhibit outcome-reporting bias involve outcomes that were initially collected but then were either incompletely reported or omitted from the final published paper entirely. One can imagine that it can be difficult for journal editors, let alone the rest of the academic community, to determine whether or not an author has omitted any outcome data prior to writing up the results and submitting the paper for publication. If the presence of outcome-reporting bias is difficult to detect one may ask the questions: How does the academic community even know about the existence of outcome-reporting bias? Or: If outcome reporting bias does exist, what would cause an author to not report certain outcomes he or she collected? The latter question is a bit more complicated, so the former will be discussed first.

Outcome-reporting bias has been shown to exist within the medical research community; a study done by Chan, Hrobjartsson, Haahr, and Altman (2004) documented such a case. In the study, the authors looked to see if researchers in the medical field in Denmark were selectively reporting/selectively withholding outcome data that were collected throughout a number of drug trials; if that were the case, outcome-reporting bias would be present. The methods they used to identify the presence of selective reporting and/or outcome-reporting bias were quite exceptional. The authors were able to uncover the protocols the original researchers created that detailed all of the outcomes that were to be collected throughout the duration of the study. When Chan et al. (2004) obtained the research protocols they were able to compare the outcomes that were listed in the
protocols to those that ended up in the original researchers’ final published paper. Upon further examination they found that a number of outcomes were described in the protocols but then never reported in the final publication. Further, they found that 71% of the outcomes that were statistically significant went on to get reported while only 56% of the non-statistically significant outcomes were reported. Failing to report any of the outcomes alone would suggest the presence of outcome-reporting bias, but this case in particular showed that more statistically significant outcomes were reported and more non-statistically significant outcomes were withheld suggesting a first potential factor associated with outcome-reporting bias: lack of statistical significance.

Chan, Krleza-Jeric, Schmid, and Altman (2004) performed a similar analysis on a number of randomized drug trials performed in Canada where the authors compared initial research protocols to final published papers. They confirmed that statistically significant outcomes were significantly more likely to be reported than non-statistically significant outcomes. In other words, outcomes that were collected but lacked statistical significance were more likely to be withheld from the final publication when compared to those that showed statistical significance. Further, the authors found that not only were outcomes lacking statistical significance withheld more readily, those outcomes that showed statistically significant negative treatment effects (harmful effects) were also withheld more readily when compared to statistically significant positive treatment effects.

Similar results were documented in American medical literature as well. Turner, Matthews, Linardatos, Tell, and Rosenthal (2008) looked at antidepressant drug trials that were documented in reviews by the Food and Drug Administration. The authors
performed a systematic review of the literature to identify subsequent publications that were made based on the data contained within these preliminary FDA reviews/drug trials.

Of the subsequent publications, the authors identified the presence of outcome-reporting bias: outcomes contained within the initial reviews done by the FDA were not reported and disseminated in the subsequent publications written by other researchers. Of all the outcomes the authors identified in the original FDA trials, 31% were not published. The authors believed that the “positiveness” of the outcome was the sole factor in determining whether or not it was reported. Of the 74 outcomes they reviewed, 37 of the 38 outcomes showing positive results were published. Conversely only 3 of the 36 outcomes that showed negative results were published while 22 of the 36 outcomes showing negative results were left forever unreported/unpublished. The authors argue that 11 of the studies showing negative results were eventually reported but done so in a way that tried to convey the negative outcome as a positive outcome. The authors found that when looking at all of the outcomes, both published and unpublished, only 51% were positive, however, when looking at only the outcomes that were reported, published, and disseminated it made it appear that 94% of the outcomes related to this antidepressant drug were positive. This dramatic change was obviously due to the presence of outcome-reporting bias. The work of Chan et al. (2004) and Turner et al. (2008) show a second potential factor associated with outcome-reporting bias: the presence of negative outcomes or potentially harmful effects associated with an intervention.

Perhaps one of the most discussed cases of outcome-reporting bias was detailed in a study done by Vedula, Bero, Scherer, and Dickersin (2009). This study also showed the selective reporting of statistically significant positive outcomes over those that lacked
statistical significance as well as those that showed statistically significant negative outcomes. In this study, the authors were able to review unpublished internal documents that circulated through the Pfizer Corporation regarding protocols for the initial drug trials of gabapentin. These unpublished documents detailed both primary and secondary outcomes that were to be collected throughout the study. The authors compared the outcomes listed in these documents to those that were eventually published. Of the 20 outcomes that were detailed in the protocols, 11 were never reported and/or published. 7 of the 9 outcomes that were reported had statistically significant positive findings mirroring a tendency to exclude data that lacked statistical significance and those that had negative effects. However, of these 7 outcomes with positive results, the primary outcome as it was defined in the published literature was different from the primary outcome defined in the protocol. It seemed that after the trials were complete the researchers changed the definitions of the primary outcomes prior to publication. The pharmaceutical company essentially promoted the drug for use in treating off-label conditions that were not substantiated by the FDA nor approved for study by the initial protocol. Vedula et al. (2009) found that the definitions of the outcomes lacking statistical significance were either changed to suggest a statistically significant positive outcome or not reported at all. When outcome-reporting bias exists in pharmaceutical research it is possible that a company could release a drug that appears to be effective when it really is not. Further, a drug could be placed on the market when certain negative or potentially harmful outcomes were never reported to the populace. It goes without saying how much of a great disservice it would be to society to imply that a drug is effective when it might not be or to withhold/hide potentially negative effects associated
with the treatment. However, pharmaceutical companies benefit financially when they release new drugs. One can assume there is a lot of pressure put on researchers to show statistically significant positive results of drug trials in order to achieve monetary gain for their respective company or funding institutions. If this apparent conflict of interest were not the case, then one may ask why the researchers for the pharmaceutical company did not publish all of their outcomes/protocols or why certain outcomes were changed prior to publication and subsequent release of the drug.

Other instances of outcome-related reporting biases are detailed in an essay written by John Ioannidis (2005) entitled: *Why Most Published Research Findings are False*. In this essay, the author claims that most of the research findings that are published in journals are based on data that has been subject to selective reporting or outcome-reporting bias as well as measurement error. An important source of bias explained by Ioannidis includes conflicts of interest in which researchers have a set agenda and attempt to bury/hide findings that are counter to that agenda. One of the ways Ioannidis identifies in which a researcher can bury findings is to suppress (or fail to report) the outcome entirely; also known as outcome-reporting bias. Some of the agendas, Ioannidis claims, that can influence researchers toward outcome-reporting bias may include personal or academic interests (a researcher’s personal or professional advancement may hinge on the success or failure of a new intervention or simply the publication of many papers) and pressures from funding agencies to deem a potentially profitable new intervention effective. Ioannidis goes further to state that the degree to which outcome-reporting bias has pervaded academic research is so severe that any attempts to replicate the results presented in published journal articles would be
practically impossible though he does state that replication is potentially the single most important guard the academic community has against bias. An article by Makel and Plucker (2014) echoed this claim made by Ioannidis by stating that only .13% of the published educational articles were replications. Further, when studies were replicated by independent researchers, similar findings to the original study could not be produced. The work of Vedula et al. (2009), Ioannidis (2005), and Makel and Plucker (2014) highlight a third potential factor associated with outcome-reporting bias: conflicts of interest when it comes to the success or failure of intervention outcomes (i.e. the pressure from funding institutions to support a new product, monetary gain, or personal/academic goals and aspirations). This factor is elaborated on even further in the work done by Simmons, Nelson, and Simonsohn (2011). In this paper, the authors show just how easy it is for researchers to manipulate data in order to produce statistically significant results and, therefore, increase the likelihood of publication. The authors argue that the goal of research should not be to publish as many articles as possible but to disseminate the truth to the academic community. Further, when researchers are put in a place where they are left with the option to either disseminate the truth or manipulate the data in order to obtain more publishable results, the researcher is almost forced to select the more self-serving option to maintain a consistent publication record. Simmons et al. (2011) posit that one way to remedy these types of choices is to increase the guidelines by which researchers and reviewers must comply prior to publication. An article by John, Loewenstein, and Prelac (2012) also supports the idea that academic misconduct happens quite often and is may be a result of questionable research practices. One of these questionable research practices identified by the authors was the exclusion of outcome
data post-hoc in order to obtain a statistically significant result. The authors issued a survey to researchers and asked them anonymously if they had ever engaged in the identified questionable research practices. Not only did John (2012) find that researchers had indeed engaged in questionable research practices, but that the researchers also tended to withhold outcome data in order to produce a statistically significant results and increase the likelihood of publication. It is clear that there is a connection between statistically significant findings, a desire to be published, and withholding outcome data.

Chan and Altman (2005) sought to examine the nature of outcome-reporting bias by reviewing published journal articles, identifying publications in which outcome-reporting bias was believed to exist, and then subsequently issuing surveys to the authors of the original publications to ask about the outcomes reported. They first identified the presence of outcome-reporting bias by seeing if the authors listed an outcome in the methods section of their papers but then did not report any results on it in the rest of the publication. Of the authors that were issued surveys, 69% of them responded. Some of the survey questions were meant to find out why certain outcomes were reported while others were not. Of the individuals who were surveyed, many of them denied that the outcomes identified as being unreported even existed despite being listed in the methods section. When Chan and Altman (2005) reviewed the surveys, they found that lack of statistical significance seemed to be the biggest reason that led to a researcher failing to report an outcome although many of the potentially harmful outcomes were also not reported. What was unique about the authors’ approach was that the issuing of surveys to researchers allowed for deeper insight into what factors may be associated with an individual failing to report an outcome that was collected. The researchers who
responded to the surveys not only highlighted lack of statistical significance as being the reason behind selective reporting, but they also noted space constraints imposed by journals and lack of clinical importance as factors leading them to intentionally omit outcomes from publications. One could ask the question: If the outcome lacked clinical importance, why was it collected in the first place? In essence, it can be assumed that lack of statistical significance probably led to this lack of reporting rather than lack of clinical importance. The more interesting of the findings from Chan and Altman (2005) highlights a fourth potential factor associated with outcome reporting bias: journal-imposed space limitations.

**Outcome-Reporting Bias in Educational Research**

The literature presented above show that the presence of outcome-reporting bias in the medical field is real and is impossible to overlook. The question then becomes what other forms of academic research show instances of outcome-reporting bias and to what degree does it pervade the literature? However, identifying the presence of outcome-reporting bias in other forms of academic research, specifically educational research, is a much more difficult task when compared to medicine (Pigott et al., 2013). In Pigott et al. (2013), the authors discuss that while in medical research the submission and publication of research protocols is commonplace, the same is not true for educational research. Many of the previously discussed studies found the presence of outcome-reporting bias by comparing the outcomes listed in research protocols to those detailed in subsequent publications. Because protocols for educational research are harder to come by it is much more difficult to ascertain exactly how many outcomes were originally collected. This fact in turn limits the ability for future researchers to determine
which, if any, of the outcomes were suppressed or selectively omitted by the researcher prior to the publication of the final paper.

Cooper, DeNeve, and Charlton (1997) noted that some institutional review boards keep records of proposals for research done in the social sciences. These proposals may include information related to what outcomes are to be collected throughout the study similar to medical research protocols. The authors claim that gaining access to these institutional review board proposal documents could provide pertinent information to aid in discovering the prevalence of outcome-reporting bias in educational research. The task would be accomplished by comparing any outcomes that were proposed but then never reported in the final publication. Although there may be some good information contained within these proposal documents, gaining access to them could be challenging.

Pigott et al. (2013) noted that many institutional review boards have differing standards and practices across institutions which could affect the accessibility of these research proposals as well as the degree to which the outcomes are completely detailed within. In addition, the authors state that the institutional review boards across universities may be comprised of individuals who are non-experts in the field of education which could result in a research proposal’s use of non-specific terminology or failure to discuss all of the outcomes in great depth. If the outcomes are not fully described in these institutional review board proposal documents there might not be enough information to discern what outcomes were collected and reported versus those that were collected and suppressed.

In order to address the apparent lack of protocols in educational research and the lack of accessibility/utility of research proposals with which to compare outcome data, Pigott et al. (2013) chose instead to compare the outcomes described in an individual’s
The authors argue that dissertations are not subject to space limitations in the same way that journal articles are and, therefore, may include information on all of the outcomes that were collected throughout the study. In addition, the authors believe that the dissertation process is freer from different forms of bias when compared to journal publications. It is possible that, if they are correct, the dissertation process could also be freer from the different factors that influence outcome-reporting bias. It is clear that doctoral dissertations are much longer and can allow for more in-depth reporting than journal publications. Because authors have more room to write and are not confined to space restrictions, there should be no reason as to why a Ph.D. candidate would not or could not include all of the outcomes he or she collected. In addition, Pigott et al. (2013) believe that the approval of academic dissertations is not subject to publication bias the way that journal articles are (the approval of a dissertation is not contingent upon the author obtaining statistically significant data). The ways in which publication bias can affect the presence of outcome-reporting bias will be discussed in greater detail later in this chapter.

The authors agreed that although it would be difficult, and potentially unhelpful, to obtain research protocols or grant proposals for educational research, looking at a candidate’s dissertation, being free from publication bias, should provide a complete record of all of the outcomes that were intended to be collected by the researcher. With a complete record of all of the collected outcomes one could easily compare the presence of these outcomes to any that were or were not reported in subsequent journal articles using the same data set. Any discrepancies between the number of outcomes collected
and reported in the dissertation and those from the subsequent journal publication would suggest the presence of outcome-reporting bias. The authors employed this method using educational dissertations and subsequent publications from 2001-2005. Ultimately, they found that non-statistically significant outcomes were 30% more likely to be omitted from a published study than statistically significant ones indicating the presence of outcome-reporting bias in the field of education. The work done by Pigott et al. (2013) provides an initial look into the presence of outcome-reporting bias in educational research. It also offers some good methods that can be used to circumvent the apparent lack of published protocols and/or access to research proposals in the field of education. Though some preliminary work was done to address the issue of outcome-reporting bias in educational research, it is still incomplete. More work needs to be done in order to fully address what factors might be associated with outcome reporting bias. This task serves as the basis for this dissertation and the process by which it can be accomplished is fully detailed in chapter three. In short, a similar strategy to the one utilized by Pigott et al. (2013) was implemented, though a greater amount information was retrieved and coded from each dissertation-publication pair with the hope of capturing the factors that may be associated with outcome-reporting bias. In addition to simply addressing the presence of outcome-reporting bias in educational research, this dissertation will detail what specific factors are associated with it as well as discuss the impact outcome-reporting bias has on the fields of education, meta-analysis, and systematic review.

Factors Influencing Outcome-Reporting Bias

In the previous section, certain factors potentially associated with outcome-reporting bias were identified through conclusions drawn from reviewing medical
literature. These potential factors were: lack of statistical significance, presence of statistically significant negative treatment outcomes or harmful effects, potential outside pressures or conflicts of interest, and lack of publication space. Each one of these factors appears to be associated with an individual researcher exhibiting outcome-reporting bias (the selective reporting or withholding of data) in their final publication. It is important to note that many of these identified factors can also be linked to other well-documented forms of bias that exist within the research community, particularly different forms of publication bias. The following section will explore how some of the identified factors above are associated with publication bias and how they may allow for outcome-reporting bias to exist. It will also show, through the use of hypothetical scenarios, how, when coupled together, publication bias and outcome-reporting bias can create a cycle that exacerbates the overall prevalence of bias in academic research.

**Publication Bias**

Lack of statistical significance was the first factor identified as may be associated with outcome-reporting bias. One of the major forms of bias which influences this factor, and subsequently outcome-reporting bias, is called publication bias. Rothstein, Sutton, and Borenstein (2005) define publication bias as any occasion in which the totality of research appearing in published literature is systematically unrepresentative of the population of completed studies. In other words, publication bias exists whenever the body of published results does not accurately reflect all of the data which were collected about a particular intervention.

The authors stress the danger in which readers are placed of drawing the wrong conclusions about an intervention because not all of the data were made available through
publication. If outcome-reporting bias were to exist within a paper and certain outcomes were withheld from publication, the academic community would not have an accurate representation of the totality of research that has been done regarding the particular intervention. This connection suggests that outcome-reporting bias and publication bias can influence one another.

Easterbrook, Berlin, Gopolin, and Matthews (1991) discuss that, from a methodological standpoint, in terms of the quality of the design and the soundness of the research, studies which produced statistically significant results do not differ very much from those which produced non-statistically significant results. In fact, the only difference between the two is that one study showed an intervention was effective and the other showed an intervention was not effective. The authors also express that both types of information are equally as important and necessary to report to potential consumers, investors, educators, etc. In other words, it is just as important for the scientific community to know that an intervention is not effective as it is to know an intervention is effective in order to make complete and accurate conclusions about its overall level of effectiveness.

Despite there being no difference in the quality of the methodological design between studies that showed statistical significance and those that did not, Dickersin, Chan, and Chalmers (1987) found that studies producing statistically significant results were three times more likely to be published than those studies that showed a non-statistically significant result. Further, the study done by Cooper, DeNeve, and Charlton (1997) showed that authors are less likely to submit papers for publication if they found
non-statistically significant results. These alarming statistics show just how the first identified factor (lack of statistical significance) is related to publication bias.

It has been established that studies lacking statistical significance are less likely to get published. Researchers are clearly aware of this fact as indicated by their apparent lack of willingness to submit papers for publication which lacked statistical significance. If authors are aware that papers lacking statistically significant data are less likely to get published, they may either choose to not submit the paper for publication or choose to withhold the outcomes that lacked statistical significance in order to convey only those outcomes that yielded statistically significant results, thereby increasing the chances for publication. For this case, consider two individuals who were separately working on demonstrating the effects of a reading intervention for English Language Learners. Both individuals used the strictest, most rigorous study design and research methodology. The only difference between the two studies was that one showed the intervention was effective (statistically significant treatment effect) and the other showed that it was not (lack of statistically significant treatment effect). Both researchers attempt to publish their results but only the individual who showed the statistically significant treatment effect gets published due to publication bias on the part of the journal editors. The educational community then believes that the treatment is effective and immediately begins its implementation despite there being evidence to the contrary. Implementing an intervention that is not actually effective could result in a substantial waste of time and money.

Consider now the same example as described above except this time the researcher who obtained the non-statistically significant results is aware that publication
bias exists but has a strong desire to get published. Instead of attempting to publish his or her data the way it is, he or she chooses to withhold some of the outcomes which suggested the intervention had no effect in favor of reporting only the outcomes that showed a significant positive effect. The researcher attempts to justify this action by claiming those data which were excluded were apparent outliers or measured incorrectly. Now this researcher’s paper demonstrates, although incorrectly, a statistically significant treatment effect, and the likelihood that it will be accepted for publication is increased. The concept of publication bias is rooted in the perception that journals do not want to publish studies that failed to produce statistically significant results and, therefore, researchers do not submit their papers for publication if they contain non-statistically significant results. If researchers are aware that journals have a bias against publishing papers that lack statistically-significant data they may choose to intentionally withhold the non-statistically significant outcomes from their work (outcome-reporting bias).

Another one of the previously identified factors potentially associated with outcome-reporting bias was whether or not the researcher obtained negative effects or harmful effects associated with the intervention. This factor is associated with a specific form of publication bias known as positive-results bias. Another consideration is that whether or not a study gets published can sometimes depend upon the researcher’s ability to demonstrate that a particular intervention is effective (publication bias). If a researcher is under a lot of pressure to get his or her research published, he or she may actively choose what to report and what to leave out in order to demonstrate that the intervention with which he or she is working is truly effective. In the Cochrane Handbook for Systematic Reviews of Interventions, Higgins & Green (2011) discuss in great detail the
concerns over a researcher’s ability to incorporate bias into their studies and the negative effects it has on the research community.

Earlier in the chapter outcome-reporting bias was said to exist when researchers collect data involving a variety of outcomes but then either incompletely report them or omit them from the final publication entirely. Another way to think about outcome-reporting bias is the selective reporting or withholding of outcome data from a publication. It is not uncommon for individuals to withhold information. Porta (2008) describes instances of individual subjects selectively reporting or withholding information when it comes to revealing different amounts of personal information such as their medical or sexual history. Porta (2008) identifies that it is completely up to the individual to decide what information he or she is willing to reveal and what information he or she is willing to suppress. When this type of bias enters research it is known as reporting bias. While Porta (2008) was referencing the selective reporting of data by subjects, it is not unreasonable to assume that researchers can do the same thing. Human beings are capable of choosing what they report and what they withhold based on personal reasons or outside influences. It is solely up to the individual to determine whether or not they choose to report accurately, falsely, incompletely, or selectively. Reporting bias and outcome-reporting bias go hand-in-hand. Individual researchers can choose what outcomes they want to reveal and what outcomes they want to withhold leading to an obvious bias within the published work. Addressing the problem of outcome-reporting bias also includes limiting the presence of publication bias as well.

By now it should be clear that outcome-reporting bias can have a tremendous effect on the academic research community as a whole. It can lead to the implementing of
interventions that were based on biased conclusions drawn from incomplete data sources. While the problem of outcome-reporting bias seems troublesome on its own, it can have a tremendous impact on the credibility of important research methodologies like meta-analysis and systematic review as well. The following section describes not only the importance of meta-analysis and systematic review for the scientific community but also how outcome-reporting bias can serve as a potential threat to the validity of any claims derived from research which incorporated these specific methodologies.

**Importance of Meta-Analysis and Systematic Review**

Outcome-reporting bias can have a significant impact on meta-analyses and systematic reviews. To better understand how outcome-reporting bias affects these important research techniques, it is probably helpful to first understand what each technique is. Although meta-analysis and systematic review are similar in nature, there are some important differences between the two processes. When an individual conducts a meta-analysis, he or she synthesizes quantitative effect sizes across multiple primary studies on a single intervention in order to calculate an average, overall effect size of said intervention. In order to compile a complete set of primary studies, a systematic review through the literature is performed (Littell, Corcoran, & Pillai, 2008). The systematic review process involves directed searching through a number of databases in order to compile a complete set of primary studies relating to the topic or intervention in question. Cooper (2010), Hedges and Olkin (1985), and Lipsey and Wilson (2001) all provide thorough looks into the steps necessary to complete a meta-analysis.

Glass (1977) discusses the importance of utilizing a technique like meta-analysis when it comes to research and data analysis in the field of education. Glass argues that
with a field like education, which is so vastly expanding, more advanced levels of analysis need to be created (Glass, 1976). A major factor guiding much of Glass’ research was that a portion of the educational research he had come across regarding the same intervention showed small sample sizes (low statistical power) and conflicting results. The weak findings derived from these data sets led the author to seek better ways to approach educational research and increase statistical power.

The use of meta-analysis by researchers is detailed in the work of John Ioannidis (2005). Ioannidis states that using a meta-analysis increases the power of a body of work compared to the power of a single study alone. The meta-analysis technique achieves this by allowing the researcher to essentially replicate each individual experiment as if they were all one large common sample by combining the effect sizes across all primary studies and calculating an average overall effect size.

In order to better conceptualize this method, consider a scenario in which five researchers are performing independent experiments on a new reading intervention across different schools. These researchers, when considered independently, may have small to moderate sample sizes due to any number of factors like lack of time, funding, or resources. Assume each researcher has sound research methods and each is able to calculate an effect size for the intervention; however, due to smaller sample sizes, their studies may lack statistical power and any results may have been affected by outliers. Each individual researcher is able to publish their findings and the educational community finds mixed results across each. Some of the effect sizes lead educators to believe the intervention was highly effective, some show only moderate to no effect. Based on these mixed findings, is it worth a school district’s time and money to work on
implementing the intervention? If all the independent researchers used sound methods, whose findings should be considered when making these important policy decisions? Was one or more of the researchers’ samples affected by outliers? This situation can be a difficult one to manage. Enter the meta-analysis. After some time and debate another researcher decides to perform a meta-analysis on the five studies in order to estimate an average, overall effect of the reading intervention. This individual combines the data across all schools effectively creating a diverse, combined sample size. With a larger sample size, not only does the study have higher statistical power when compared to each of the original five studies independently, but it also is able to estimate the average, overall effect of the intervention. Any effect of outliers or bias in the previous studies is minimized and the educational community can have a much better understanding regarding the true nature of the effectiveness of the reading intervention and can make their policy decisions accordingly. However, despite the apparent utility and importance of the meta-analysis and systematic review techniques, the problem of outcome-reporting bias can serve as a potential threat to the validity of any findings derived from these types of studies.

Potential Threats to Validity

If any primary studies within the educational literature have excluded certain outcomes from publication, the findings from those studies are biased. Similarly the findings from any subsequent meta-analysis or systematic review performed using those primary studies would also be biased (Orwin & Cordray, 1985). The problem of outcome-reporting bias can definitely affect the validity of claims made by meta-analyses and systematic reviews and, subsequently, the overall perception of these two important
techniques as viable research options. In order to ensure that meta-analyses and systematic reviews are taken seriously and that any findings derived from such work are valid, the academic community must understand the prevalence of outcome-reporting bias and what factors influence it.

The problem of outcome-reporting bias is real and has been documented in several disciplines including education, though much of the work that has already been done regarding outcome-reporting bias comes from the medical research community. That being said, a review of the medical literature has allowed for several factors potentially associated with the presence of outcome-reporting bias to emerge. It is the purpose of this dissertation to discover if these same factors are associated with the presence of outcome-reporting bias in the field of education. By identifying what factors are associated with outcome-reporting bias in educational research, the academic community can begin taking steps to reduce, and hopefully eliminate, those factors. By ensuring that educational research includes all of the outcomes that were collected on an intervention not only increases the validity of the claims, but it also increases the validity of subsequent meta-analyses and systematic reviews performed on the intervention as well.
CHAPTER THREE

METHODS

The methods utilized for this dissertation were built upon those described by Pigott et al. (2013). This chapter will not only provide an overview of the methods described by Pigott et al. (2013), it will also explain how those methods were replicated and updated for this dissertation project in order to help answer the research questions:

1. Do differences exist between the outcomes reported in an individual’s doctoral dissertation and those reported in subsequent publications utilizing the same data set?
2. If outcome-reporting bias exists in the identified sample of dissertations and published articles, what factors are associated with its presence?
   a. Are statistically significant outcomes more likely to get reported than those that are not statistically significant?
   b. Are statistically significant positive outcomes more likely to get reported than those that showed statistically significant negative outcomes or harmful effects?
   c. Are outcomes consistent with the author’s initial hypothesis more likely to get reported than those do not?
   d. Do authors having pursued academic careers in which publication is necessary have higher instances of outcome-reporting bias than those that did not?
   e. Are there any factors associated with sample demographics that lead to higher instances of outcome-reporting bias?
Established Methods of Comparison

In the medical field, outcome-reporting bias is identified by comparing outcomes listed in research protocols to those reported in the final publications. If any of the outcomes are withheld from publication, outcome-reporting bias is present. However, research protocols are not necessarily commonplace in educational research. One way outcome-reporting bias can be identified within educational literature is by comparing the outcomes reported in an individual’s dissertation to those that were reported in subsequent journal publications. Making this type of comparison is based on the assumption that journal articles are much shorter in length and more prone to other types of bias when compared to dissertations; therefore, certain information, including outcomes, within a dissertation may get omitted from subsequent journal publications indicating outcome-reporting bias.

The major task became how to properly locate educational dissertations that listed intervention outcomes and also had subsequent journal publications associated with those same outcomes. Pigott et al. (2013) hypothesized that individuals who obtained doctoral degrees from institutions having very high research activity were more likely to pursue academic careers that required frequent publication. Using this assumption one would be able to select institutions that had very high research activity and search for dissertations approved by those institutions. Once a complete list of dissertations is compiled, a thorough search among a variety of research databases would reveal if any of the doctoral candidates made any subsequent journal publications associated with the same
intervention outcome data. A more thorough description of the ways in which this process was utilized will be detailed later in this chapter.

**Selecting Research Institutions in the Original Study**

The Carnegie Classification system was used in the original publication by Pigott et al. (2013) in order to identify institutions that have very high research activity. The work of McCormick and Zhao (2005) detailed 96 research institutions that were given a designation of “RU/VH” by the Carnegie Classification system indicating that these 96 universities had “very high” research activity. Since the work of McCormick and Zhao (2005) the Carnegie Classification system was updated to include a total of 108 universities with the “RU/VH” designation in 2010. The system will be updated again early in 2015. The work of Pigott et al. (2013) used the original 96 research institutions described in 2005. Because one of the purposes of this dissertation is to build upon the original data set compiled by Pigott et al. (2013), the same 96 institutions were identified and utilized. It is important to note that any future research into the presence of outcome-reporting bias in educational research should update the search to include those institutions newly listed as having very high research activity by the Carnegie classification system.

**Search Strategies in the Original Study**

The work of Pigott et al. (2013) utilized ProQuest Dissertations and Theses database to search for all Ph.D. and Ed.D. dissertations which were completed among the 96 institutions with very high research activity between the years of 2001 and 2005. The word “Education” was used a keyword and “Ph.D. or Ed.D.” was used to specify degree earned within the database. This process was completed for all 96 “RU/VH” universities
and returned over 6,000 potential dissertations. All dissertations returned underwent an abstract review to determine whether or not they list any educational intervention outcomes. Any dissertations that were deemed as not having data relating to educational intervention outcomes were not included in the subsequent analyses. In addition, abstracts were screened to include only those interventions for pre-K to 12th grade students in order to include only the outcomes associated with the grades in which students are required to be in school and were collected in an educational setting. Any observational studies or qualitative designs were excluded as primary outcomes for these types of dissertations might not be as clear to identify. Because most of the medical research on outcome-reporting bias included only experimental or quasi-experimental interventions (drug trials), the medical literature served as the basis for identifying potential factors associated with the presence of outcome-reporting bias, and primary outcomes are more difficult to identify in qualitative/observational studies, the use of only experimental and quasi-experimental designs became the primary focus for the original study. After applying the aforementioned criteria to the 6,000 dissertations, Pigott et al. (2013) ended up with 607 potential dissertations to include. The next step in the process performed by Pigott et al. (2013) became identifying subsequent articles that matched the data described in these 607 dissertations.

**Identifying Subsequent Articles in the Original Study**

The search process for this stage began with utilizing academic journal search engines such as Google Scholar, PsychINFO, and ERIC. The subsequent articles were identified by applying search criteria within the previously mentioned databases that include combinations of the authors’ names, the titles of the dissertations, and certain
keywords found within each dissertation. Once a subsequent article had been identified, it was paired with the original dissertation in order to make direct comparisons between each. In any event where it became unclear whether or not an article should be paired with a particular dissertation, a deeper look was taken into the actual data sets contained within each to ensure they utilized the same data sets. Any dissertation that failed to get paired with a journal article was excluded from further analysis. The authors ended up with a total of 79 dissertation-article pairs. Once all dissertations were paired with subsequent journal articles, the coding process began. This step involved extracting a larger amount of information from each dissertation/article pair. For the purposes of this dissertation, it became important to replicate these methods described by Pigott et al. (2013), in order to make sure no dissertation/article pairs were missed and to serve as a reliability check to the original work.

**Reliability Checks and Replicated Searches**

Orwin and Vevea (2009) recommend that multiple researchers perform the same searching and coding processes independently in order to establish reliability within a body of academic work. Pigott et al. (2013) had identified a total of 79 dissertation/article pairs for further analysis from an initial search that returned over 6,000 dissertations. The first task in this dissertation became independently replicating the search for articles to pair with 607 dissertations that met the initial search criteria as described by Pigott et al. (2013). This step was used to make sure no subsequent article was missed and no dissertation was erroneously excluded from further analysis when it should have been included due to a failure to identify an appropriate article pair. Most importantly, this step served as a means to check the reliability of the original paper. For this project, the names,
titles, and keywords from the original 607 dissertations identified in the study by Pigott et al. (2013) were used as search terms among Google Scholar, PsychINFO, and ERIC in order to identify any subsequent articles that were published utilizing the same data sets. For this study, some additional dissertation/article pairs were identified. The original study identified 79 pairs while this new search returned 83 pairs. It is important to note that these 83 dissertation/article pairs included all 79 of the original pairs identified by Pigott et al. (2013). Therefore, this dissertation included further analysis on 83 dissertation/article pairs rather than 79 as described in the original study. With these 83 dissertation/article pairs identified, the coding process began.

**Coding and Data Analysis**

The coding process for this dissertation involved pulling several pieces of information from each article/dissertation pair. Some of the information that was pulled from each dissertation and article included: the author’s name, the title, and the year of publication. These pieces of information were included to make identification of the dissertation/article pairs easier. The other types of information that were coded were done so to specifically address the research questions.

In order to address the first research question (do differences exist between the outcomes reported in an individual’s doctoral dissertation and those reported in subsequent publications utilizing the same data set?) the outcomes that were listed in each dissertation were identified, numbered and recorded. By comparing the total number of outcomes present in a dissertation to those that were reported in the paired publication, one could see if any outcomes were indeed withheld. An example of the coding manual that was used in the answering of this research question is presented in the table below:
Table 1. Coding Information for Study and Outcome Identification

<table>
<thead>
<tr>
<th>Question:</th>
<th>Code:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study ID</td>
<td>Coder’s first initial, Coder’s last initial, # (BT1, BT2, etc.)</td>
</tr>
<tr>
<td>Dissertation Author</td>
<td>Dissertation author’s last name, author’s first name</td>
</tr>
<tr>
<td>Dissertation Title</td>
<td>Title of the dissertation</td>
</tr>
<tr>
<td>Dissertation Year</td>
<td>Year the dissertation was published</td>
</tr>
<tr>
<td>Publication Author</td>
<td>Publication author’s last name, author’s first name</td>
</tr>
<tr>
<td>Publication Year</td>
<td>Year the journal article was published</td>
</tr>
<tr>
<td>Number of Outcome</td>
<td>Beginning with 1 for each dissertation, number each outcome</td>
</tr>
<tr>
<td>Name of Outcome</td>
<td>Briefly describe the nature of the outcome collected</td>
</tr>
</tbody>
</table>

This coding instrument was used, in part, for questions 2a-2e.

For research question 2a (are statistically significant outcomes more likely to get reported than those that are not statistically significant?) information regarding the level of statistical significance of each outcome and whether or not the outcome was reported was recorded among each dissertation/article pair. The information that was coded to help in the analysis of this question is presented in the table below:

Table 2. Coding Information Regarding Statistical Significance and Reporting

<table>
<thead>
<tr>
<th>Question:</th>
<th>Code:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Used</td>
<td>Name of the statistical test used to measure the outcome</td>
</tr>
<tr>
<td>Sample Size</td>
<td>Sample size for the outcome</td>
</tr>
<tr>
<td>Sample Size df</td>
<td>Degrees of freedom for the sample</td>
</tr>
<tr>
<td>Alpha Used</td>
<td>Alpha-level associated with the outcome</td>
</tr>
<tr>
<td>p-value</td>
<td>p-value associated with the test statistic for the outcome</td>
</tr>
<tr>
<td>Test Statistic</td>
<td>Type of test statistic used</td>
</tr>
<tr>
<td>Test Statistic Value</td>
<td>Value of the test statistic</td>
</tr>
<tr>
<td>Test Statistic df 1</td>
<td>Degrees of freedom associated with the test statistic</td>
</tr>
<tr>
<td>Test Statistic df 2</td>
<td>Degrees of freedom associated with the test statistic</td>
</tr>
<tr>
<td>Significant?</td>
<td>0 = Test statistic was not significant 1 = Test statistic was significant</td>
</tr>
<tr>
<td>Reported?</td>
<td>0 = The outcome was not present in the journal article 1 = The outcome was present in the journal article</td>
</tr>
</tbody>
</table>

This coding instrument was used, in part, for questions 2a-2e.
For each outcome present in the dissertations, the information in the above table was recorded. An outcome was coded with a “1” for “Reported” if the outcome within the dissertation was also found in the subsequent publication; similarly, the outcome was coded “0” if the outcome was present in the dissertation but not present within the later publication.

In order to address question 2b (are statistically significant positive outcomes more likely to get reported than those that showed statistically significant negative outcomes or harmful effects?) it first became important to operationally define a “statistically significant negative outcome” or a “harmful effect.” Many academic interventions are designed and implemented in order to increase positive outcomes or decrease negative outcomes. A statistically significant “negative outcome” or “harmful effect” was identified and coded if the intervention had a statistically significant effect in the opposite direction for which the intervention was intended. In other words, if an intervention were designed to increase the number of words a student was able to type per minute but the outcome showed that the intervention actually decreased the number of words a student was able to type per minute, that outcome was coded as a “harmful effect”. Similarly, if an intervention were intended to decrease the number of detentions a student had over the course of a semester but the student ended up significantly increasing the number of detentions they had over the course of a semester, a “harmful effect” was also coded. The table below shows the coding manual for this question:

<table>
<thead>
<tr>
<th>Question:</th>
<th>Code:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harmful</td>
<td>0 = A harmful effect was not attributed to this outcome</td>
</tr>
<tr>
<td></td>
<td>1 = A harmful effect was attributed to this outcome</td>
</tr>
</tbody>
</table>

This coding instrument was used, in part, for question 2b.
For question 2c (are outcomes consistent with the author’s initial hypothesis more likely to get reported than those that do not?) the significance level and the direction of the author’s hypothesis and the conclusion were noted. In other words, if an author expected an intervention to work equally well for boys and girls, but girls performed significantly better and boys showed no effect or a significant decrease, that outcome was coded as “not consistent with the author’s hypothesis.” Similarly, if an author expected an intervention to have a significant increase in students’ reading abilities, but the intervention showed no effect, the outcome was also coded as “not consistent with the author’s hypothesis.” These definitions show that question 2c is indeed different from asking only questions about significance or only about harmful effects. The coding manual used for this question is presented below:

Table 4. Coding Information for Consistency

<table>
<thead>
<tr>
<th>Question:</th>
<th>Code:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistent</td>
<td>0 = Outcome was not consistent with the author’s hypothesis</td>
</tr>
<tr>
<td></td>
<td>1 = Outcome was consistent with the author’s hypothesis</td>
</tr>
</tbody>
</table>

This coding instrument was used, in part, for question 2c.

Addressing question 2d (do authors having pursued academic careers in which publication is necessary have higher instances of outcome-reporting bias than those that did not?) was a little more difficult. For this question, a number of demographic data were collected including whether or not the author’s held a faculty position at a university when their subsequent articles were published and of the academic institution for which they worked. Because this question is meant to capture potential publication bias and conflicts of interest, the authors’ jobs and affiliations were noted when available to see if individual’s having pursued faculty careers exhibit more frequent instances of outcome-reporting bias when compared to researches not in professorship roles. This was
accomplished by doing a subgroup analysis within the larger dataset. One could parse out only the individuals that had faculty positions at the time of journal publications and look at the degree to which outcome-reporting bias might have been present. This finding could then be compared to a similar subgroup analysis in which only the individuals that held positions outside of faculty roles were considered. The affiliations of the author’s were ascertained through a search through the article to find the job title of the author at the time of publication. The data that needed to be coded for this question are presented below:

Table 5. Coding Information for Academic Affiliations

<table>
<thead>
<tr>
<th>Question:</th>
<th>Code:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author’s Affiliation</td>
<td>$0 = $Author did not have a faculty position when published $</td>
</tr>
<tr>
<td></td>
<td>$1 = $Author did have a faculty position when published $</td>
</tr>
</tbody>
</table>

This coding instrument was used, in part, for question 2d.

For the final research question (are there any factors associated with sample demographics that lead to higher instances of outcome-reporting bias) a number of sample characteristics were also pulled from each dissertation/article pair. Information like primary race/ethnicity, socio-economic status, grade level, and sex of the sample were recorded when available. These pieces of information were pulled in order to assess whether or not research done on any particular groups might be more subject to outcome-reporting bias. These questions were also answered via a variety of subgroup analyses.

For a complete look at what information was pulled from each dissertation/article pair, please see the attached coding manual in the appendices. This coding manual is intended to serve as a basis for others hoping to perform future research in this area and should include the newly designated “RU/VH” institutions as they become available or other identified factors. Trikalinos and Ioannidis (2005) advocate all coding manuals be
well-defined and easily accessible as future work will always need to be done in order to adequately update systematic reviews and meta-analyses. The reliability of the coding manual was checked by having two researchers independently code the first five dissertation/article pairs. Once this step was completed the researchers discussed any discrepancies that arose until a consensus was reached and any necessary changes to the coding manual were made. Once inter-rater reliability had been established, the rest of the dissertation/article pairs were coded by the primary investigator.

For each dissertation/article pair, a weighted average odds-ratio was estimated using a Mantel-Haenszel meta-analytic approach (Shadish& Haddock, 2009). This process was used because there were a varying number of statistical tests that were conducted within each dissertation/article pair. Because each dissertation/article pair had a different number of statistical tests and, therefore, outcomes, it would not be appropriate to sum across all cells as in the traditional mean odds-ratio approach. This Mantel-Haenszel odds-ratio can be calculated by the following equation:

\[
\text{OR}_{\text{MH}} = \frac{\sum_{i} a_i d_i}{\sum_{i} n_i} \times \frac{\sum_{i} b_i c_i}{\sum_{i} n_i}
\]  

(EQ 1)

Where \( i = 1 \ldots k \) for every dissertation/article pair and (a), (b), (c), and (d) correspond to the number of outcomes present within each dissertation/article pair that met certain criteria. For example, in order to address the question regarding statistical significance, (a) became a measure of the number of statistically significant outcomes within each pair, i, that were reported; (b) became a measure of statistically significant outcomes that were not reported in each pair; (c) became the number of non-statistically significant outcomes
that were reported for each pair; and (d) was used to represent the number of outcomes that were not statistically significant and were not reported for each pair. This type of analysis strategy served as the basis for analyzing the remaining research questions where the definitions for a, b, c, d, and e changed depending on the specific question asked. For some of the research questions it became necessary to make comparisons between subgroups of the 83 dissertation/article pairs. An example of this technique can be seen in the analysis of question 2d in which the Mantel-Haenszel odds-ratio was calculated for both faculty members and non-faculty members to see which subgroup showed higher instances of outcome-reporting bias. Examples of the ways in which contingency tables were created in order to be used in equation (1) are presented below:

Table 6. Data Collected From Each Dissertation/Article Pair For Use in Question 2a.

<table>
<thead>
<tr>
<th></th>
<th>Reported in Published Paper</th>
<th>Not Reported in Published Paper</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistically Significant</td>
<td>a&lt;sub&gt;i&lt;/sub&gt;</td>
<td>b&lt;sub&gt;i&lt;/sub&gt;</td>
<td>a&lt;sub&gt;i&lt;/sub&gt; + b&lt;sub&gt;i&lt;/sub&gt;</td>
</tr>
<tr>
<td>Not Statistically Significant</td>
<td>c&lt;sub&gt;i&lt;/sub&gt;</td>
<td>d&lt;sub&gt;i&lt;/sub&gt;</td>
<td>c&lt;sub&gt;i&lt;/sub&gt; + d&lt;sub&gt;i&lt;/sub&gt;</td>
</tr>
<tr>
<td>Totals</td>
<td>a&lt;sub&gt;i&lt;/sub&gt; + c&lt;sub&gt;i&lt;/sub&gt;</td>
<td>b&lt;sub&gt;i&lt;/sub&gt; + d&lt;sub&gt;i&lt;/sub&gt;</td>
<td>n&lt;sub&gt;i&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

*Note:* The number of statistically significant outcomes that were also published in a subsequent article were reported in a<sub>i</sub> then used in equation (1).

Table 7. Data Collected From Each Dissertation/Article Pair For Use in Question 2b.

<table>
<thead>
<tr>
<th></th>
<th>Reported in Published Paper</th>
<th>Not Reported in Published Paper</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistically Significant Positive Effect</td>
<td>a&lt;sub&gt;i&lt;/sub&gt;</td>
<td>b&lt;sub&gt;i&lt;/sub&gt;</td>
<td>a&lt;sub&gt;i&lt;/sub&gt; + b&lt;sub&gt;i&lt;/sub&gt;</td>
</tr>
<tr>
<td>Statistically Significant Negative or Harmful Effect</td>
<td>c&lt;sub&gt;i&lt;/sub&gt;</td>
<td>d&lt;sub&gt;i&lt;/sub&gt;</td>
<td>c&lt;sub&gt;i&lt;/sub&gt; + d&lt;sub&gt;i&lt;/sub&gt;</td>
</tr>
<tr>
<td>Totals</td>
<td>a&lt;sub&gt;i&lt;/sub&gt; + c&lt;sub&gt;i&lt;/sub&gt;</td>
<td>b&lt;sub&gt;i&lt;/sub&gt; + d&lt;sub&gt;i&lt;/sub&gt;</td>
<td>n&lt;sub&gt;i&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

*Note:* The number of statistically significant outcomes that showed a positive effect and were also published in a subsequent article were reported in a<sub>i</sub> then used in equation (1).
Table 8. Data Collected From Each Dissertation/Article Pair For Use in Question 2c.

<table>
<thead>
<tr>
<th></th>
<th>Reported in Published Paper</th>
<th>Not Reported in Published Paper</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result Consistent with Author’s Initial Hypothesis</td>
<td>$a_i$</td>
<td>$b_i$</td>
<td>$a_i + b_i$</td>
</tr>
<tr>
<td>Result Not Consistent with Author’s Initial Hypothesis</td>
<td>$c_i$</td>
<td>$d_i$</td>
<td>$c_i + d_i$</td>
</tr>
<tr>
<td>Totals</td>
<td>$a_i + c_i$</td>
<td>$b_i + d_i$</td>
<td>$n_i$</td>
</tr>
</tbody>
</table>

*Note:* The number outcomes that were consistent with the author’s initial hypothesis and were also published in a subsequent article were reported in $a_i$ then used in equation (1).

Table 9a. Data Collected From Each Dissertation/Article Pair For Use in Question 2d.

<table>
<thead>
<tr>
<th></th>
<th>Reported in Published Paper</th>
<th>Not Reported in Published Paper</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faculty Position &amp; Statistically Significant</td>
<td>$a_i$</td>
<td>$b_i$</td>
<td>$a_i + b_i$</td>
</tr>
<tr>
<td>Faculty Position Held &amp; Not statistically Significant</td>
<td>$c_i$</td>
<td>$d_i$</td>
<td>$c_i + d_i$</td>
</tr>
<tr>
<td>Totals</td>
<td>$a_i + c_i$</td>
<td>$b_i + d_i$</td>
<td>$n_i$</td>
</tr>
</tbody>
</table>

*Note:* The number of outcomes that were presented by an individual holding a faculty position and were also published in a subsequent article were reported in $a_i$ then used in equation (1).

Table 9b. Data Collected From Each Dissertation/Article Pair For Use in Question 2d.

<table>
<thead>
<tr>
<th></th>
<th>Reported in Published Paper</th>
<th>Not Reported in Published Paper</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Faculty Position &amp; Statistically Significant</td>
<td>$a_i$</td>
<td>$b_i$</td>
<td>$a_i + b_i$</td>
</tr>
<tr>
<td>Non-Faculty Position Held &amp; Not statistically Significant</td>
<td>$c_i$</td>
<td>$d_i$</td>
<td>$c_i + d_i$</td>
</tr>
<tr>
<td>Totals</td>
<td>$a_i + c_i$</td>
<td>$b_i + d_i$</td>
<td>$n_i$</td>
</tr>
</tbody>
</table>

*Note:* The number of outcomes that were presented by an individual holding a faculty position and were also published in a subsequent article were reported in $a_i$ then used in equation (1).

The same technique was used in Table 4a and Table 4b to make subgroup comparisons between a number of sample characteristics. This technique was used to assess whether or not there were any groups to which researchers showed increased
levels of outcome-reporting bias. For this question, a subgroup comparison was made between predominantly white and predominantly non-white samples, high school and non-high school samples, special education and general education samples.

Pigott et al. (2013) suggest that the best way to quantify the results of these Mantel-Haenszel odds-ratios is to convert each one into a risk-ratio. By doing this conversion, the results can be better interpreted as a “percent more likely to be withheld from publication.” The risk-ratio calculation can be expressed as follows:

\[
RR = \frac{\sum_{i} a_i(c_i + d_i)}{\sum_{i} n_i} \quad \text{EQ 2}
\]

\[
RR = \frac{\sum_{i} c_i(a_i + b_i)}{\sum_{i} n_i}
\]

Where \( i = 1 \ldots \ldots k \) for every dissertation/article pair.
CHAPTER FOUR

RESULTS

The original paper by Pigott et al. (2013) identified 79 dissertation/article pairs for their analysis. After replicating the search protocol, the same 79 dissertation/article pairs were identified with the addition of 4 others bringing the total to 83 dissertation/article pairs. Within these 83 dissertation/article pairs there were a total of 1,599 outcomes for which the statistical significance (or lack thereof) and information regarding whether or not the outcome was later published could be ascertained. As mentioned in the previous chapter, a variety of information was extracted and coded from each pair in order to better answer the research questions presented below:

1. Do differences exist between the outcomes reported in an individual’s doctoral dissertation and those reported in subsequent publications utilizing the same data set?

2. If outcome-reporting bias exists in the identified sample of dissertations and published articles, what factors are associated with its presence?
   a. Are statistically significant outcomes more likely to get reported than those that are not statistically significant?
   b. Are statistically significant positive outcomes more likely to get reported than those that showed statistically significant negative outcomes or harmful effects?
   c. Are outcomes consistent with the author’s initial hypothesis more likely to get reported than those do not?
d. Do authors having pursued academic careers in which publication is necessary have higher instances of outcome-reporting bias than those that did not?

e. Are there any factors associated with sample demographics that lead to higher instances of outcome-reporting bias?

The table below presents some of the initial descriptive statistics from the dissertation/article pairs:

Table 10. Descriptive Statistics for Dissertation/Article Pairs

<table>
<thead>
<tr>
<th></th>
<th>Number of Outcomes</th>
<th>Number of Statistically Significant Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissertations</td>
<td>1,599</td>
<td>651</td>
</tr>
<tr>
<td>Articles</td>
<td>803</td>
<td>371</td>
</tr>
</tbody>
</table>

N = 83 Dissertation/article pairs.

**Do Differences Exist Between the Dissertations and the Articles?**

In order to answer the first research question (do differences exist between the outcomes reported in an individual’s doctoral dissertation and those reported in subsequent publications utilizing the same data set?) one would simply need to look at the total number of outcomes listed within the dissertations and compare it to the total number of outcomes reported within the subsequent publications. Among the 83 dissertations, there were a total of 1,599 outcomes listed. Among the 83 publications, there were 803 outcomes reported. To be more specific: given the set of 83 dissertations containing 1,599 total outcomes, only 803 were later reported in subsequent journal articles. This finding means that only about 50.2% of the total outcomes that were collected were eventually reported in later publications and 49.8% were not reported at all. This large discrepancy answers the first research question that differences do indeed
exist between the outcomes reported in dissertations and those reported in subsequent publications using the same data set.

Because these discrepancies exist, one may wonder why educational researchers are not reporting all of the outcomes that they collected for their dissertation projects. A number of factors potentially linked to outcome-reporting bias have been identified in chapter two of this dissertation by way of looking at cases of documented outcome-reporting bias in the medical field. One of the major factors identified that may be linked to outcome reporting bias had to do with the idea of statistical significance.

<table>
<thead>
<tr>
<th>Table 11. Reporting Statistically Significant Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Outcomes</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>Dissertations</td>
</tr>
<tr>
<td>Articles</td>
</tr>
</tbody>
</table>

N = 83 Dissertation/article pairs.

Of the 1,599 outcomes listed within the dissertations, about 651 of them were statistically significant and about 948 were not statistically significant. To summarize the data even further, within the 76 dissertations, about 40.7% of the outcomes were statistically significant and 59.3% lacked statistical significance. If there were no preference for educational researchers to selectively report or selectively withhold outcome data with regard to statistical significance (or lack thereof) a similar ratio of statistically significant to non-statistically significant outcomes should be contained within the 76 subsequent articles when compared to the dissertations. However, when looking at the 76 articles alone, 371 outcomes were statistically significant (46.2%) and
were not statistically significant (53.8%). The change in these ratios when looking at the types of data that are reported between dissertations and journal articles does not appear to be that drastic but when one considers only the outcomes that were excluded (withheld) from publication, 64.8% of them were not statistically significant. So there was an apparent difference that existed between the outcomes listed in a dissertation and those reported in a subsequent journal article. Is this difference big enough to assume that educational researchers are intentionally withholding data based solely on its lack of statistical significance? In order to answer this question, a more sophisticated approach was needed.

**Outcome-Reporting Bias and Statistical Significance**

The first factor that could be related to outcome-reporting bias worth considering is statistical significance. Specifically, are statistically significant outcomes more likely to get reported than those that are not statistically significant? As mentioned previously, a more sophisticated approach was used to answer this research question; in this case, a Mantel-Haenszel odds-ratio was calculated. The Mantel-Haenszel odds-ratio was used because not all of the 83 dissertation/article pairs had the same number of statistical tests contained within. This method allows one to control for the differing number of statistical tests that were used within each dissertation/article pair. The Mantel-Haenszel odds-ratio across the 83 dissertation/article pairs was about 2.41 (with a 95% confidence interval ranging from a lower limit of 1.84 to an upper limit of 2.98). As mentioned previously, a more appropriate analysis technique to employ would be to calculate a risk ratio associated with these values. The risk-ratio came out to be 1.30. This risk-ratio indicates
that outcomes failing to show statistical significance are 30% more likely to be withheld from publication.

**The Effect of “Positiveness”**

Because statistically significant outcomes appeared to be published and non-statistically significant outcomes appeared to be withheld more regularly, the question then became whether or not it was the mere statistical significance that was related to outcome-reporting bias or if there was something more to consider. Are statistically significant negative results or harmful effects published or withheld with the same discretion as statistically significant positive results? In order to answer this question a similar approach was taken to that described in the previous section. Each dissertation/article pair was coded for the degree of “positiveness” of the outcome as well as whether or not the outcome was reported. The “positiveness” of an outcome was determined by whether or not any significant negative or harmful effects associated with an intervention were present. These “harmful” effects were defined as any outcome in which the intervention had a statistically significant effect that was directly opposed to the primary intent of the intervention. Consider another brief example: if a claim were to be made that a particular reading intervention should significantly increase the number of words a student is able to read within a time frame but in reality the intervention significantly decreased the number of words a student read within that timeframe, it would be coded as a statistically significant negative/harmful effect. Similarly, if a claim were to be made that an intervention would decrease the number of swear words a student uses within a given time frame but the intervention showed a statistically significant increase in swear words used over that time frame, that outcome would be
coded as a statistically significant negative/harmful effect as well. Out of the 76 dissertation/article pairs that were coded for this dissertation, there were no outcomes that reported any type of negative or harmful effects associated with a treatment or intervention. For a deeper consideration of these results, please see the next chapter. Because none of the outcomes had statistically significant negative/harmful effects, further analysis could not be completed.

**Results Consistent with the Authors’ Hypotheses**

Statistical significance appeared to play a role in outcome-reporting bias. It became interesting to see if educational researchers also chose to selectively report/withhold outcomes that were not consistent with their original hypotheses. Perhaps an individual researcher thought that an intervention would have no effect; or perhaps they believed the intervention outcomes would not be different between two different demographic groups. What happens when one considers presence or absence of outcome-reporting bias aside from whether or not the outcome was statistically significant and instead considers whether or not the researcher “accurately predicted” what would happen? Consider an example in which a researcher is testing a new math intervention. This researcher believes it will help all individuals and believes that there will be no difference when it comes to socioeconomic status (i.e. all SES levels will have the same amount of gains given this intervention). The researcher implements the intervention and not only do all students improve their math scores, the intervention itself does not appear to be biased toward higher or lower socioeconomic statuses. In this case, both the researcher’s hypotheses are correct, the intervention worked, and showed no statistical difference between SES groups. Outcomes like these would be coded as “consistent with
the author’s original hypothesis.” Given the outcomes from the 76 dissertation/article pairs, the Mantel-Haenszel odds-ratio came out to be 2.91 (with a 95% confidence interval range including an upper limit of 3.48 and a lower limit of 2.34). Looking at the associated risk-ratio (1.41), it would appear that outcomes that are not consistent with the authors’ original hypotheses are 41% more likely to get withheld from publication than those that are consistent. This finding suggests that educational researchers are more likely to report outcomes in which they could correctly predict the outcome and were more likely to suppress outcomes about which their hypothesis was “wrong.” The table below shows some descriptive statistics relating to outcomes and whether or not they were consistent with the authors’ initial hypotheses.

Table 12. Results Consistent with the Authors’ Hypotheses

<table>
<thead>
<tr>
<th></th>
<th>Number of Outcomes</th>
<th>Number of Outcomes Consistent with the Authors’ Hypotheses</th>
<th>Number of Outcomes Not Consistent with the Authors’ Hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissertations</td>
<td>1,599</td>
<td>633</td>
<td>966</td>
</tr>
<tr>
<td>Articles</td>
<td>803</td>
<td>353</td>
<td>450</td>
</tr>
</tbody>
</table>

N = 83 Dissertation/article pairs.

Out of all the 1,599 outcomes, 633 were consistent with the original hypothesis put forth by the authors (39.6%). Of those outcomes that were consistent, 353 ended up getting reported (55.8%) while 280 (44.2%) were suppressed. This finding means that of the 1,599 outcomes present within this data set, 966 were not consistent with the hypotheses of the authors (60.4%). Of these non-consistent outcomes, only 450 ended up getting reported (46.6%) while 53.4% of the outcomes that were not consistent remained unreported. These numbers would definitely suggest a preference authors have toward reporting outcomes that are consistent with their own hypotheses.
Academic Affiliations and Outcome-Reporting Bias

John Ioannidis (2005) suggested that an individual’s career in academia may play a role into their selective inclusion or omission of outcome data in final published reports. The idea behind this belief is that researchers who have faculty positions are under a much greater pressure to publish papers because tenure decisions are (partially) based upon an individual demonstrating the ability to continually publish new research. Because publication bias has been documented across many disciplines (i.e. journals are more likely to publish papers that have shown statistical significance), these faculty members may be biased toward only submitting articles for publication that have demonstrated statistical significance of an intervention. Because not all interventions show statistical significance, it is possible that an individual may selectively include statistically significant outcomes while selectively excluding outcomes that failed to show statistical significance in their final published reports. In order to test this idea, the authors’ academic affiliations were coded from each journal article. If the author had held a faculty position at the time of the final publication it was coded as a “1” while if the author had held a non-faculty position (high school teacher, principal, nature camp supervisor, etc.) it was coded as a “0” for comparison. If it could not be ascertained as to whether or not an individual had held a faculty position at the time of publication the article was not included in further analysis. Of the 83 dissertation/article pairs, 63 authors had held faculty positions at the time of publication (75.9%) while only 20 (24.1%) held jobs other than faculty positions.

For this analysis, 1,380 of the outcomes came from dissertation/article pairs that belonged to individuals holding faculty positions at the time of publication while 219 of
the outcomes came from individuals who did not hold faculty positions at the time of publication.

For the individuals holding faculty positions, 41.4% of the outcomes were statistically significant while 58.6% were not. For these individuals, about 50.1% of the outcomes were reported while 49.9% of the outcomes were suppressed. For the individuals not holding faculty positions, 36.5% of the outcomes were statistically significant while 63.5% were not. The interesting thing to note is that the individuals who did not have faculty positions published 52.5% of their outcomes while suppressing 47.5%; a very similar comparison to their faculty counterparts.

Table 13. Academic Affiliations Subgroup Analysis

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Number of Outcomes</th>
<th>Odds-Ratio</th>
<th>Confidence Interval</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faculty</td>
<td>1,380</td>
<td>2.19</td>
<td>1.6-2.76</td>
<td>1.26</td>
</tr>
<tr>
<td>Non-faculty</td>
<td>219</td>
<td>3.97</td>
<td>3.4-4.54</td>
<td>1.50</td>
</tr>
</tbody>
</table>

N = 83 Dissertation/article pairs.

To continue looking into this issue, the Mantel-Haenszel odds-ratio was calculated for just the individuals holding faculty positions and then again for just the individuals who did not hold faculty positions. For the individuals who held faculty positions at the time of publication, the odds-ratio came out to be around 2.19 (with a 95% confidence interval ranging from a lower limit of 1.62 to an upper limit of 2.76). The risk-ratio for the faculty members was 1.26 indicating that outcomes failing to show statistical significant were 26% more likely to get suppressed from the final publications than those that were statistically significant. For the individuals not holding faculty positions, the odds-ratio came out to be 3.97 (with a 95% confidence interval ranging from a lower limit of 3.4 to an upper limit of 4.54). The risk-ratio for non-faculty
members was 1.50 indicating that outcomes failing to show statistical significance were 50% more likely to appear in final publications than those that were not statistically significant.

**Other Potential Factors Identified**

A review of the literature had shown that the factors discussed above could potentially be associated with outcome-reporting bias. It is possible that this list is incomplete. In other words, just because a potential factor was identified through a review of the literature does not mean that there are not other factors that might be associated with outcome-reporting bias which have yet to be documented. Further coding was completed in order to see if any other sample-level characteristics could speak to the nature of outcome-reporting bias in educational research. Some of these characteristics included: the grade level of the students in the intervention, the primary race/ethnicity of the samples, the primary SES of the samples, whether or not the study focused on special-education populations, and the setting of the intervention (rural, urban, suburban). For many of these factors, further analysis was not completed for several reasons. For example, only a few of the dissertation/article pairs focused on special education populations and many of the pairs focused on mixed settings (used students from rural, suburban, and urban settings combined). These potential factors could still be associated with outcome-reporting bias but too few dissertation/article pairs had diverse enough characteristics to make any further analysis meaningful.

Two comparisons that were able to be made was between samples that focused on high school populations compared to those that focused on elementary and middle school
populations and those that focused on predominantly white populations compared to those that did not.

Table 14. Grade Level Subgroup Analysis

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Number of Studies</th>
<th>Odds-Ratio</th>
<th>Confidence Interval</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School</td>
<td>31</td>
<td>2.31</td>
<td>1.74-2.88</td>
<td>1.25</td>
</tr>
<tr>
<td>Non-HS</td>
<td>52</td>
<td>2.45</td>
<td>1.88-3.02</td>
<td>1.32</td>
</tr>
</tbody>
</table>

N = 83 Dissertation/article pairs.

The Mantel-Haenszel odds-ratio for high school populations was 2.31 (with a 95% confidence interval of 1.74-2.88) with a risk-ratio of 1.25. These results indicate that outcomes failing to show statistical significance are 25% more likely to be excluded from publication than those that showed statistical significance among high school samples.

The results were similar for elementary/middle school samples. The Mantel-Haenszel odds-ratio was 2.45 (95% confidence interval 1.88-3.02) with a risk ratio of 1.32. These results indicate that outcomes failing to show statistical significance are 32% more likely to be excluded from publication than those that showed statistical significance among elementary and middle school samples.

Table 15. Race Subgroup Analysis

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Number of Studies</th>
<th>Odds-Ratio</th>
<th>Confidence Interval</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predominantly White</td>
<td>65</td>
<td>2.07</td>
<td>1.53-2.61</td>
<td>1.24</td>
</tr>
<tr>
<td>Non-White Samples</td>
<td>18</td>
<td>6.84</td>
<td>6.27-7.41</td>
<td>1.73</td>
</tr>
</tbody>
</table>

N = 83 Dissertation/article pairs.

Among predominantly white samples, the Mantel-Haenszel odds-ratio was 2.07 (95% confidence interval 1.53-2.61) with a risk-ratio of 1.24. These results indicate that outcomes failing to show statistical significance are 24% more likely to be excluded from
publication than those that showed statistical significance among predominantly white samples. Further, The Mantel-Haenszel odds-ratio for samples of minority (non-white) predominance was 6.84 (95% confidence interval 6.27-7.41) with a risk ratio of 1.73. These results indicate that outcomes failing to show statistical significance are 73% more likely to be excluded from publication than those that showed statistical significance among non-white samples.
CHAPTER FIVE
DISCUSSION

This dissertation was meant to provide insight into the nature of outcome-reporting bias in educational research. The specific goal of this project was to identify factors that may be associated with outcome-reporting bias in order to suggest ways in which educational researchers and journal editors alike can improve their personal and professional practices. The factors presented within this dissertation were identified by looking at instances in which outcome-reporting bias was documented in other academic fields and hypothesizing what factors might have been associated with those particular instances of said bias. Once these factors were identified, a systematic search among educational dissertations and articles was performed in order to see if any of the factors were present within them. The previous chapter outlined some of the major findings from this process. This section will provide a summary of those results, discuss the implications of them, and suggest ways in which the academic community can use these results to improve practices among educational researcher and journal editors in order to ensure that any information that is disseminated to the academic community is based on sound practices and valid and reliable data.

Differences that Exist Between Dissertation/Article Pairs

The previous chapter showed that not all outcomes which are collected by educational researchers for their dissertations end up getting published in subsequent journal articles. This fact begs the question: Are there reasons as to why an author would
exclude certain outcomes from the final publication? Chan and Altman (2005) asked that very question to a variety of researchers in the social sciences. The major reason that authors gave for excluding outcomes from journal publications was that most journals impose a maximum space limitation. This space limitation requires researchers to self-edit out data prior to getting published. The question then becomes: If authors have to selectively edit or exclude information from their papers prior to publication, how do they decide which information to keep, and which information to eliminate? It is evident from looking at the results that outcomes are indeed eliminated from the final publications. This systematic exclusion of outcome data suggests that there may be a bias when it comes toward the reporting of outcomes. Therefore, it is clear there is indeed a difference in the outcomes reported in an individual’s doctoral dissertation and those reported in subsequent publications utilizing the same data set. Out of the 1,599 outcomes present within the sample of dissertations, only 50.2%, that is 803 outcomes, were later reported in a journal article. It is definitely clear that authors are showing some type of bias when it comes to the selective inclusion/exclusion of outcome data when it comes to journal publications.

If not all outcomes are made available to the academic community then how are informed decisions supposed to be made with regard to interventions? What type of outcome information is being withheld? The answer to questions such as these really speaks to the nature of outcome-reporting bias in educational research. Further, this problem only compounds the “file drawer problem” present in the fields of meta-analysis and systematic review. An individual hoping to perform one of these important methodologies on a given intervention would not have access to all of the information
that was collected from the research. This problem could lead to the validity of the findings derived from meta-analysis and systematic review to be called into question.

Because authors identified space limitations as a primary reason leading them to exclude data from their articles, increased publication space allotments could potentially be one solution to this problem of outcome exclusion. However, it would be difficult for the academic research community to do away with space limitations entirely due to the increased publishing costs that would be involved with allowing researchers unlimited submission space. That being said, with the advent of primarily online journal databases and repositories, space limitations may be a little more flexible than they once were.

Perhaps the best possible solution to the problem that arise due to differences that exist between the number of outcomes that were collected and those that were reported is to encourage researchers to be more forthright with their data. Authors who have had to exclude certain portions of data from their publications could include information within their papers indicating to readers that more outcomes were collected but excluded do to space limitations. These researchers could include endnotes, hyperlinks, e-mail addresses, etc. within their papers in order to inform subsequent readers about the existence of other outcome data that were collected throughout the course of the study and provide them with a path by which to obtain access to the full set of outcome data. Full transparency should be the expectation and standard, especially when it comes to interventions that could potentially help our students. Including information within a publication regarding the existence of other outcomes that were not included in the final publication should be considered best practice. Another option would be for the researcher to break their study up into several articles so that all of their outcome data could be disseminated to the
research community. However, because of publication bias, there is no guarantee that these subsequent articles would get published. It is for this reason that I believe including some type of explanation in the first publication should become standard.

**Outcome-Reporting Bias and Statistical Significance**

It has been documented that publication bias exists. Journals are less likely to publish papers that lack statistical significance and researchers are less likely to submit papers for publication if they contain outcomes that lack statistical significance. Because differences exist between the outcomes that were collected in an individual’s dissertation and those that were reported in subsequent publications, one can assume that presence or lack of statistical significance could be a major factor associated with the selective inclusion or exclusion of outcome data. In other words, if authors are aware that lack of statistical significance decreases their likelihood of being published, they may choose to selectively exclude the non-statistically significant outcomes from publication while including only those outcomes that showed statistical significance.

The data presented in Chapter Four illustrated that statistically significant outcome data were more likely to get published than those outcomes that lacked statistical significance. In fact, non-statistically significant outcome data were 30% more likely to get excluded from publications when compared to the outcomes that demonstrated statistical significance. This result clearly shows that authors are indeed systematically excluding certain types of outcomes from publication. Publication bias appears to play a role in this particular factor, but why is there a tendency for journals to only publish statistically significant findings? There appears to be a misconception present in the academic community that lack of statistical significance indicates that the
research is not worth publishing because it did not show anything particularly remarkable or new or innovative. Maintaining beliefs such as these does a great disservice to the academic community. It is just as important for the academic community to know when interventions have no effect as it is to know when they have a statistically significant impact.

Picture a school district spending much time, money, and other resources in order to implement an educational intervention. In the end, the intervention shows no effect. This failed implementation not only caused the taxpayers a lot of money, but it also wasted the time and energy the school district could have put into finding another intervention that might have been successful. Now imagine that research was done showing that intervention was not effective but it was not published due to a publication bias. Had this research been made available, the school district could have saved much time and money.

Consider another example. In this scenario a researcher collects outcomes on an intervention and shows no statistical significance. Believing that they will not get published unless they show statistical significance, they selectively withhold/hide/explain away the outcomes that were not statistically significant and only attempt to publish the outcomes showing statistical significance (outcome-reporting bias). Not only does the school district still lose time and money by implementing an intervention that is not effective, but any researcher hoping to perform a meta-analysis or systematic review using this data set is now working with invalid data and biased findings.

These examples show just how important it is for the academic community to have access to all outcome data that were collected. It is only when quality practices are
used and full transparency is given that truly informed decisions can be made. A potential solution to the problem associated with the selective exclusion of statistically significant outcome data would be to encourage editors of journals to do away with these preconceived beliefs that data failing to produce significant differences are not interesting or important. Achieving this goal requires a complete change in the culture of research. Publication bias with regard to statistical significance is so ingrained in researchers’ minds that outcome-reporting bias becomes necessary if one wants to get published when they have data showing no statistical significance. Perhaps increasing the access to and use of online publication databases is the answer to this problem as well. With a seemingly infinite amount of space to store data, these databases could include all publishable work including both papers that showed statistically significant outcomes and those that did not show statistically significant outcomes. Perhaps a new online journal or online database needs to be created in which all peer-reviewed academic papers can be published regardless of the statistical significance of their outcomes. This would allow everybody access to a full spectrum of studies.

The Effect of “Positiveness”

The idea that harmful outcome associated with interventions might be selectively withheld from publication was identified as a potential factor through a review of the medical literature. However, in exploring this case when it comes to education, not a single dissertation reported negative or harmful effects associated with an intervention or treatment. At first this situation appeared suspicious because there were plenty of examples in which medical treatments produced harmful effects in patients but not many when it came to education. One can arrive at two conclusions based on the fact that not a
single dissertation listed a statistically significant treatment outcome. The first is that individual authors withhold statistically significant treatment outcomes even at the level of their dissertations and that is why no harmful treatment effects were present. The other more likely conclusion is that by simply implementing an intervention designed to improve an academic outcome, students will either improve or stay the same. In other words, if a school district were to implement an intervention to improve math scores, by simply attempting to address the issue, students will either achieve at a greater level or remain at the same level as if no intervention were implemented at all. The next step in analyzing this question for the field of education would include incorporating a wider search net to see if any academic interventions produced statistically significant harmful effects associated with treatment.

**Results Consistent with the Authors’ Hypotheses**

Although there appeared to be a tendency for authors to withhold outcomes that failed to show statistical significance while going on to publish statistically significant findings, many non-statistically significant outcomes were indeed published. The question was then asked whether publishing non-statistically significant outcomes was done so in order to present as many findings to the academic community as possible or if authors were simply publishing the non-statistically significant outcomes because they were consistent with what the author originally believed to be true. Thus, it became important to explore whether or not having an outcome consistent with a hypothesis had any relationship to selective reporting. Ultimately, regardless of statistical significance, results that were not consistent with the authors’ original hypotheses were 41% more likely to get suppressed than those that were consistent. This statistic shows that authors
clearly have a tendency to publish findings that are consistent with their own personal claims rather than trying to publish as many outcomes as possible. If there were no preference, the results would have shown that authors publish results that are both consistent and not consistent with their original hypotheses with a bit more equity. This issue can create a major problem in the research field because it appears that authors are not being forthright with their data and only trying to verify their own claims. This type of bias raises a huge red flag when it comes to the validity of educational research as well as findings derived from meta-analyses and systematic reviews based on educational data sets. Without access to all of the relevant information and collected outcomes, the validity of any claims in the educational research field is highly questionable. Stressing the importance of transparency and completeness should be paramount when it comes to educating would-be researchers. Emphasizing good practices and research ethics should be paramount. A paradigm shift needs to take place where being “right” is not valued more than being thorough. Personal agendas need to be put aside in favor of validity.

**Academic Affiliations and Outcome-Reporting Bias**

Perhaps one of the best places in which a paradigm shift can take place is when it comes to research institutions pressuring their faculty and non-faculty members alike to continually publish papers. When looking at only individuals who held faculty positions at the time of publication, non-statistically significant outcomes were 26% more likely to get suppressed than statistically significant outcomes. This figure is a stark comparison to their non-faculty counterparts where non-statistically significant outcomes were 50% more likely to get suppressed. These data suggest that publication bias, and, subsequently, outcome-reporting bias, may affect researchers holding both faculty and non-faculty
positions with the latter being at greater risk. Faculty members are under a great deal of pressure to continually publish papers and justify their own jobs. Because it has been well documented that publication bias exists in favor of statistically significant outcomes, faculty members may resort to selectively withholding non-statistically significant outcomes in order to increase the likelihood that their papers will get accepted and published. The interesting piece was that non-faculty members were more likely to omit outcomes from publications. This fact would suggest that either pressure to publish exists with greater prevalence in non-academic settings or there is some other outside factor leading non-faculty members to feel the need to suppress outcome data that lacks statistical significance. These problems again causes a great deal of concern when it comes to the results that are presented to the academic research community as well as the findings derived from meta-analyses and systematic reviews based on previous work. The only way to correct this issue is to encourage better practices among researchers and publishers alike. Lack of statistical significance does not imply that the work is not important or necessary to be disseminated to the public. Increasing access to non-statistically significant research outcomes and minimizing the bias that researchers and publishers have toward statistically significant outcomes is a necessity.

Other Potential Factors, Final Thoughts, and Suggestions for Future Research

Some of the other potential factors that are worth taking another look at involve sample characteristics. Outcomes generated from non-white samples as well as elementary schools and middle schools appear to be much more subject to outcome-reporting bias. These findings suggest certain areas in which the findings derived from meta-analyses and systematic reviews might be more subject to bias and, therefore, less
valid. It is important that all outcomes be made available, particularly among these identified sample groups so as to allow more valid claims to be disseminated to the research community.

Although a few factors were identified in this dissertation, the exploration into the nature of outcome-reporting bias is far from over. There may be many other reasons authors are selectively excluding outcome data from their final publications. The work presented here should serve as a basis upon which future research can be done to better understand the problem that is outcome-reporting bias. It is clear that outcome-reporting bias is a major problem in the field of educational research and needs to be dealt with accordingly. The data that are being presented to the academic community are based on biased findings and incomplete data sets. Researchers in the field of meta-analysis are at a great disadvantage because their important methodology relies on having access to all outcome data. Meta-analysis is a research tool that will only gain momentum in the coming years and it is vital that the academic community can trust the findings of any meta-analysis are valid. This can only be done if individuals are encouraged to be completely transparent with their outcome data. The best way to ensure this type of transparency is to encourage researchers and publishers to move away from their own personal biases against certain types of data and to remove any type of stigma or pressure associated with a researcher presenting “incorrect” or “un-publishable” data. Whether this is achieved by educating students at the college level or by adopting rigorous research standards, something needs to be done to combat the problem of outcome-reporting bias in educational research.
The creation and increased use of online databases can only serve to help correct these problems. Online databases are freer from space limitations and can house publications that include non-statistically significant research outcomes. Databases such as these are also seemingly infinite and eternally searchable which helps increase the accessibility of all sources of data that were collected for all types of studies. Having online databases that are easily searchable also helps individuals hoping to perform meta-analyses find relevant studies.

Hopefully the work presented within this dissertation can raise awareness of outcome-reporting bias and provide some insight into factors that are potentially related to it. Finally, this dissertation is meant to serve as a call to action for everyone in the educational research community to work together to combat the problem that is outcome-reporting bias and hopefully, someday soon, eliminate it from practice entirely. It is only when biases like outcome-reporting bias are eliminated that the academic research community can trust that educational research conclusions are valid and, most importantly, truly effective interventions can be implemented in order to better serve our students.
APPENDIX A

CODING INSTRUMENT
Does the study focus on academic outcomes?  
Is the study experimental or quasi-experimental?  
Does the study relate to an academic intervention?  
Does the study take place in a pre-K – 12 environment?

Question:

Study ID  
Dissertation Author  
Dissertation Title  
Dissertation Year  
Dissertation Institutions  
Publication Author  
Publication Year  
Publication Author’s Affiliation  
Publication Source  
Outcome (start with 1 for each dissertation/article pair)  
Name of Outcome  
Test Used  
Sample Size  
Sample Size df  
Alpha  
p-value  
Test Statistic  
Test Statistic Value  
Test Statistic df 1  
Test Statistic df 2  
Significant?  
Published?  
Consistent With Author’s Hypothesis?  
Negative/Harmful Effects Associated with Outcome?  
Level of Study  
Type of Design  
Pretest Differences  
Personal Data Set?  
Grade Level of Sample  
Longitudinal?  

Code

First Initial, Last Initial, #  
Last name, Initials  
Title of Dissertation  
Year  
School that conferred degree  
Last name, Initials  
Year  
Where was the article obtained?  
# of outcome listed in dissertation  
Describe the outcome collected  
Name of statistical test used to measure the outcome  
Sample size for outcome  
Degrees of freedom for sample  
Alpha used  
p-value used  
Type of test statistic used  
Value of calculated test statistic  
Degrees of freedom associated with test statistic  
Degrees of freedom associated with test statistic  
0 = No, 1 = Yes  
0 = No, 1 = Yes  
0 = No, 1 = Yes  
1 = Student level, 2 = School level, 3 = District level, 4 = State level, 5 = Country level  
0 = Quasi, 1 = Experiment  
0 = No pretest differences  
1 = Pretest differences identified  
1 = Personal data set 0 = Large set  
0 = PreK-K, 1 = 1st grade, etc.  
0 = No, 1 = Yes
REFERENCES


VITA

Brian Trainor was born and raised in Melrose Park, Illinois. He graduated from West Leyden High School in 2003. After high school, Brian attended Illinois Wesleyan University, where he received a Bachelor of Science in biology; and National-Louis University, where he received a Master of Arts in education. Prior to attending Loyola, Brian worked as a high school teacher in the state of Illinois, teaching classes in all levels of math and science.

While at Loyola, Brian was awarded a graduate assistantship which covered his tuition and promoted his scholarship throughout the course of his dissertation process. Brian’s tenure at Loyola allowed him to participate as a statistician and consulting research methodologist on a number of projects both with faculty and students at Loyola as well as faculty members at other institutions throughout the United States. Brian also serves as an active member of the American Educational Research Association (AERA) and has presented papers at their national conference.

Brian was given the unique opportunity to work as an adjunct faculty member at Loyola, teaching undergraduate classes in statistics. He also was able to participate in designing and co-teaching a graduate-level course in Causal Inference Models for the Graduate School in the Research Methodology program.