

POTENTIAL JURORS' PERCEPTIONS OF POLYGRAPHS IN COURT

by

Jacob Schiess

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Arts in Criminal Justice

Boise State University

May 2018

Jacob Schiess
SOME RIGHTS RESERVED



This work is licensed under the Creative Commons Attribution- 4.0
International License. To view a copy of this license, visit
<http://creativecommons.org/licenses/by/4.0/> or send a letter to
Creative Commons
PO Box 1866
Mountain View, CA 94042

BOISE STATE UNIVERSITY GRADUATE COLLEGE

**DEFENSE COMMITTEE AND FINAL READING
APPROVALS**

of the thesis submitted by

Jacob Schiess

Thesis Title: Potential Jurors' Perceptions of Polygraphs in Court

Date of Final Oral Examination: 9 March 2018

The following individuals read and discussed the thesis submitted by student Jacob Schiess, and they evaluated the student's presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Lisa G. Bostaph, Ph.D. Chair, Supervisory Committee

Charles R. Honts, Ph.D. Member, Supervisory Committee

Laura King, Ph.D. Member, Supervisory Committee

The final reading approval of the thesis was granted by Lisa Growette Bostaph, Ph.D., Chair of the Supervisory Committee. The thesis was approved by the Graduate College.

DEDICATION

For my truly wonderful mother, whose kindness and belief in me pushes me to aspire to greatness. You have been my greatest supporter in life and I always hope to make you proud. Additionally, I thank God for granting me the benefit of a very dear friend whose companionship and understanding were very much needed through this process. I wouldn't have come through this nearly as well as I have without you, Sa'ad – thank you. I am also grateful for Melissa, who was a kind voice in a time of life where one was very much needed.

ACKNOWLEDGEMENTS

Many thanks to Dr. Charles Honts for making the literature review process much easier by providing a huge Dropbox folder full of sources. Thanks to Dr. Andrew Giacomazzi without whom this idea may never have occurred to me at all.

ABSTRACT

The polygraph occupies a contentious place in the justice system. The instrument detects various physical responses and records these results, and the examiner interprets the readings and makes a determination on whether the test subject was truthful or deceptive. Polygraphs are, in some jurisdictions, a part of the court process and in others are forbidden. On the whole, there is less research on the polygraph and their permissibility in the legal process compared to other types of evidence. There is even less research on the opinions of jurors, especially compared to surveys of criminal justice professionals. That which is present is inconsistent. This thesis was intended to measure the opinions of a pool of potential jurors on the relative weight and veracity they assign to the polygraph. With the noted inconsistency, this research was primarily exploratory and replicative in nature. To obtain data, a 17-question online survey was administered to students in nine selected courses. Professors in these courses either forwarded the survey link to their students via email or posted the link on Blackboard. It was emphasized that the survey was strictly voluntary. There were three hypotheses: respondents would have only moderate faith in the polygraph, criminal justice students would have harsher views than those not in such courses, and that those selected to receive the extra literature summary would have harsher views than those who did not. Results of the study only substantiated the first hypothesis. Chi-square analysis showed an almost complete lack of significance in theorized relationships. Receipt of the additional literature summary only

affected the respondents' general opinions of polygraph evidence and was insignificant for every other dependent variable.

TABLE OF CONTENTS

DEDICATION	iv
ACKNOWLEDGEMENTS	v
ABSTRACT	vi
LIST OF TABLES	x
LIST OF ABBREVIATIONS	xi
CHAPTER ONE: INTRODUCTION.....	1
CHAPTER TWO: LITERATURE REVIEW	4
Polygraph Science and Theory	4
Types of Polygraph Exams	7
Polygraph Use in Court.....	9
Polygraph Law and Court Challenges	11
Support for Polygraphs	16
Opposition to Polygraphs.....	19
Perceptions of the Justice System	40
Perceptions of Crime and Sentencing	44
Perceptions of Courts	49
Citizen Opinion on Polygraphs in Comparison	51
Validity of Mock Trial / Jury Research	55
CHAPTER THREE: METHODOLOGY	60

Research Questions and Hypotheses	61
Conceptual Definitions	62
Study Design, Sampling Procedure, and Data Collection	63
Dependent Variables.....	67
Independent Variables	68
Validity and Reliability.....	70
Limitations	72
CHAPTER FOUR: RESULTS	75
Sample Descriptives.....	75
Bivariate Analyses	78
Discussion.....	84
Suggestions for Future Research	88
Conclusion	90
REFERENCES	91
APPENDIX A.....	98
APPENDIX B	100
APPENDIX C	109
APPENDIX D.....	112
APPENDIX E	114

LIST OF TABLES

Table 1.0	Sample descriptives	76
Table 1.1	Support for examiner explanation of non-deceptive results	79
Table 1.2	Support for examiner explanation of deceptive results.....	79
Table 1.3	General opinion of the polygraph	79
Table 1.4	Perceived accuracy of the polygraph	80
Table 1.5	Chi-square analyses CJ course with dependent variables.....	81
Table 1.6	Bivariate correlations	82
Table 1.7	Chi-square analyses literature summary with dependent variables	83
Table 1.8	Bivariate correlations for literature summary	84

LIST OF ABBREVIATIONS

BSU	Boise State University
CQT	Control Question Test
GKT	Guilty Knowledge Test
ICE	Institute for Credentialing Excellence
DODPI	Department of Defense Polygraph Institute (now ICE)
APA	American Polygraph Association
ASTM	American Society of Testing and Materials
FRE	Federal Rules of Evidence
MRE	Military Rules of Evidence
DNA	Deoxyribonucleic Acid
ANOVA	Analysis of Variance
AAPOR	American Association for Public Opinion Research
IRB	Institutional Review Board

CHAPTER ONE: INTRODUCTION

This thesis examined the opinions of potential jurors on the scientific validity and continued necessity of polygraphs in the criminal justice system. The science appears solid, since most polygraph instruments test for involuntary physiological responses to questions (National Research Council, 2003). A baseline response is typically drawn from questions that people should have no issue passing, such as their name. When the body reacts differently than this baseline, the examiner will ask further questions regarding the matter, possibly concluding with their professional judgment that the person is being deceptive (National Research Council, 2003). Approximately 80 years ago, the polygraph made its official debut in the courtroom judging guilt in a murder trial (*Wisconsin v. Grignano & Loniello*, 1935).

The primary research question addressed was: what is the level of belief that the average college student (as a potential juror) has in polygraph use and accuracy? If this could be answered definitively, then it would lead to additional questions: what degree of use should the polygraph have? In a more general sense, does the public trust the polygraph? This research was intended to be a descriptive study of college students, as formulating policy implications would require a larger, more publicly representative sample than was achieved here.

However, the practice and theory of polygraphing is not exact, nor are the results always accurate (Steinbrook, 1992). Even earlier than the murder trial mentioned above, the U.S. Supreme Court had determined in *Frye v. United States* (1923) that the

instrument had not gained scientific acceptance in the community and therefore was inadmissible in courts. An example of this lack of acceptance can be observed on an episode of the well-known television show *Mythbusters*, where the build team tested several methods that purportedly allowed one to beat the polygraph, that is, to be deceitful and go undetected. While two out of their three attempts failed, one succeeded, which was evidence that it was possible to lie on a polygraph. Criticism of the polygraph frequently is rooted in disagreement over its accuracy. According to detractors, polygraph instruments are not the only source of error; examiners and the methods of the test are also to blame. As a whole, critics argue that the polygraph is simply too fallible to have a role in the justice system (Gallai, 1999).

For polygraph evidence to be admitted into courts, it must pass the *Daubert* test. The *Daubert* test requires the party seeking polygraph introduction to demonstrate to the judge's satisfaction its validity, replicability, and scientific support for the practice. To pass this test, it is helpful for polygraph proponents to be able to refer to general polygraph theory. Honts and Quick (1995) clarified what exactly is meant by the term 'theory'. Honts and Quick's (1995) stated definition is appropriate here: a theory is not a statement of fact, but scientific statements of belief in causal relationships between two or more variables (Honts & Quick, 1995). While a statement of unequivocal fact was practically impossible when dealing with the social sciences and especially human subjects, it was unclear whether the results in the literature on polygraphs met the second part of the stated definition for theory. While literature certainly did exist that supports the polygraph, virtually all of it only suggested a relationship, not causality. Polygraph advocates often emphasized the theoretical value of the instrument within the justice

system: that it can provide evidence of crime-relevant knowledge even in the face of a suspect who is not forthcoming. Federal and state agencies also continue to improve the standards of instrument and examiner, as well as emphasize the importance of using technology to aid in examiner analysis. Progress has certainly been made in these fields, and some jurisdictions have begun to allow the polygraph in court (*Lee v. Martinez, 2004*).

Unfortunately, in comparison to such topics as recidivism, probation, and attitudes toward police, there is a comparative dearth of information on polygraphs. In particular, there is a notable shortage of studies that analyze the degree of trust the public has in polygraphs. One issue that unfortunately could not be rectified is that results of most polygraph research were classified by the U.S. government, which made them inaccessible for reference. This study primarily identified the opinions on comparative polygraph veracity, with a minor component asking respondents to grade their own ability to detect lies informally. The remainder of the thesis includes a review of present literature on the topic of polygraphs, the methodology utilized in the study, and a discussion of its weaknesses and strengths. The final sections address the results, analyses, and the conclusions drawn from the data.

CHAPTER TWO: LITERATURE REVIEW

To reach a conclusion in the judicial part of the legal process, evidence is necessary. To convict an individual in a criminal court, the standard of proof required is ‘beyond a reasonable doubt’ (*Woolmington v. DPP*, 1935). Evidence can be of many types, though the most important in the context of this research is that of polygraph examinations. The primary research question addressed was: what is the level of belief that potential jurors have in the polygraph’s accuracy and usage in courtroom decisions? Secondly, are potential jurors educated enough to distinguish between the differing types of polygraph evidence: positive detection against negative non-detection? Lastly, the study also briefly examined the perceptions of personal lie detection: how accurate do individuals believe they are at identifying lies informally?

Polygraph Science and Theory

The theoretical value of the polygraph in the courtroom is significant. Because court decisions involve determinations of fact and fiction, a way to objectively establish fact would be of tremendous practical value. The prosecution and defense (as well as their involved witnesses) have an interest in ensuring that the result is favorable to them. Due to the presence of this interest, it is important that the justice system have a way to establish or investigate the objectivity of the evidence that these individuals provide. Ibek (2012) summarized it concisely, explaining that in a courtroom, a polygraph could be used to demonstrate that a person has crime-relevant knowledge despite outward denial.

When speaking of polygraphs, it is very important to understand that there is not one particular way to conduct all exams (National Research Council, 2003). Not only can the instrument vary in what it measures (e.g., some gauge stress responses from the skin – i.e. sweat, rushes of blood, others may monitor heart rate / pulse), but the way that questions are formed in the interview and phrased on the exam will vary by person, agency, examiner, and format of test (National Research Council, 2003). Honts and Perry (1992) nonetheless emphasized that these differences in measurement do not always amount to a statistically significant result on the polygraph itself. It is also important to be clear on the three types of results of polygraph exams: non-deceptive, inconclusive, and deceptive (Honts, 2013). Non-deceptive results indicate that the examiner and instrument found no deception. Deceptive results, as one might expect, indicate that deception was detected. Inconclusive results indicate that a judgment could not be made on whether deception was present (Honts, 2013).

Despite this, there are still common themes to all polygraphs (National Research Council, 2003). Every investigative polygraph will involve a pre-exam interview, where the examiner will talk to the subject at length. There are multiple reasons for this interview. First, it allows the examiner to explain the test procedure to the individual in detail as well as the possibility to address any concerns the test subject may have (National Research Council, 2003). Second, the interview also gives the examiner a reasonable length of time to establish an understanding of the emotional state of the subject. Third, the interview can involve a brief demonstration of the polygraph's function where the subject may be asked to lie about an unrelated topic, and the examiner

will show that the instrument is capable of detecting the deception (National Research Council, 2003).

While neural imaging is separate from polygraphs, understanding the science of lying is important to polygraphs as well. Neural imaging is a relatively new development, where magnetic resonance imaging (MRI) is used to identify and measure brain activity. This technology can and has been used to study the brain's autonomic response to lying. One of the assumptions of guilty knowledge test (GKT) polygraphs is that lying places more stress on the brain than telling the truth. GKTs are a format of polygraph exam that attempt to determine whether the test subject possesses the 'guilty knowledge' for which the test is named. Neurons may respond almost instantly to a stimulus, but the corresponding changes in blood flow to certain areas of the brain have a delay of one to two seconds, on average (Greely & Illes, 2007). Some polygraphs and imaging exams use blood pressure as indicators of activity, and while blood flow is a reasonably consistent measure for activity, this connection is not perfect (National Research Council, 2003). Because the brain also controls several autonomic functions that sustain life, there is not a true zero point for comparison (Greely & Illes, 2007). Greely and Illes (2007) clarified that muscle movement and corresponding blood flow also occur naturally during a polygraph or investigation, as individuals shift and engage in whatever small behaviors they have to in order to feel comfortable. One of the problems that blood pressure-based polygraphs encounter is individuality: each person, especially those with disabilities or traumatic injuries, will respond differently. This not only makes interpreting the test results of one individual more difficult, but makes it even more problematic to try to establish any standard, since no response would be repeatable (Greely & Illes, 2007). A

standard is generally preferable over individual interpretations, because as Raskin and Kircher (2014) indicated, a lack of standard is a risk factor for bias.

Types of Polygraph Exams

Iacono (2008) also explained that one method of polygraph administration, the relevant-irrelevant test, is not used in investigative manners because of its potential intimidation factor. A relevant-irrelevant polygraph exam compares the responses of two types of questions. Relevant questions are obviously those which are pertinent to the crime at hand (Iacono, 2008). An example of this form of question, in a case of homicide, could be, "Did you shoot the victim in the head?". An irrelevant question is one that is designed to provoke no noticeable response and is frequently an obviously-true statement such as, "Is the sky blue?". The problems associated with this form of polygraph exam are quickly evident. It is vague and in the opinion of the National Research Council (2003) should not be used. The relevant-irrelevant test in particular has a high risk of eliciting false positives (Iacono, 2008). The physiological fear response associated with being criminally detected and of a falsely positive detection is not always distinguishable for a polygraph instrument and can be interpreted as criminal in nature by the examiner (Iacono, 2008). Nevertheless, Iacono (2008) did not call for an abolition of polygraph use, explaining that it has value in investigations, provided the results can be corroborated. In some instances, corroboration can be achieved by administering another type of polygraph.

The control question test (CQT) format is similar in some respects to the relevant-irrelevant test, though with some improvements. In this format, there are two types of questions. The relevant questions are defined identically as in the relevant-irrelevant test;

those which inquire about matters relevant specifically to the subject of investigation (National Research Council, 2003). Control questions in the CQT ask of generally undesirable acts, which can but does not always include behaviors similar to the topic of investigation. In a battery case, a control question might be, “Have you ever fought anyone?”, where a relevant question would be, “Did you fight John Smith on this date?” (National Research Council, 2003). The theory behind CQT tests is that innocent individuals will respond more strongly towards the comparison questions because the questions may still deal with criminal behaviors, while the guilty will respond strongly to the relevant questions because they have specific knowledge or memory of the crime being investigated (National Research Council, 2003).

Staunton and Hammond (2011) examined the use of what practitioners call the guilty knowledge test (GKT), how physical responses can vary based on gender, and how different polygraph types measure their responses for deception. The GKT process includes examiner questions which have one relevant choice, typically a criminal detail, and control alternatives which would be inconsequential to those with no knowledge of the matter. Ben-Shakhar, Bar-Hillel, and Kremnitzer (2002) explained that the theoretical basis of the GKT is in the orienting responses of humans. To understand this, a simple definition is required. An orienting response is the brain’s subconscious reaction to a change or stimulus in an environment, but not enough to merit the startle response (Ben-Shakhar et al., 2002). Individuals who have the ‘guilty knowledge’ the exam is testing for will unconsciously react when pertinent questions arise regarding it, even if they try to suppress it. Previous analyses have only focused on the positive influence of measures based on the electric responses in the skin, which is the most common method utilized.

Staunton and Hammond (2011) found that this emphasis leads to an incomplete and sometimes erroneous picture of guilt or innocence. They concurred that lying is cognitively demanding and that lying does tend to result in arousal of the autonomic nervous system, which is the overall premise of polygraphs. This study was similar in this respect to the research of Staunton and Hammond (2011), as it did not intend to challenge the premise of polygraphs. It is worth mentioning that Honts and Perry (1992) found that GKT polygraphs were considered, among federal law enforcement, to be of very little worth, to the point where they were not even discussed in a typical polygraph course.

Polygraph Use in Court

Faigman, Fienberg and Stern (2003) explained the discretionary license given to polygraph examiners and the overall lack of clarity of the purpose of such tests. They also provided a useful historical narrative which details the expansion of polygraphs and overall historical trends. Beginning with the *Frye* case, polygraphs immediately ran into issues when brought into the courtroom. For most of the next 70 years, courts typically prohibited the introduction of polygraphs as evidence on the basis of them failing to meet the ‘general acceptance’ standards set forth by *Frye* (Faigman et al., 2003). *Daubert* modified these standards, avoiding the usage of the vague term ‘general acceptance’ as a standard. Faigman et al. (2003) began their analysis with the contention that the polygraph has overstepped its bounds. Faigman et al. (2003) argued that no scientific justifications have been discovered for the polygraph’s employment, and that it was overall based on an unfounded supposition. They continued on to claim that the polygraph falls under the scope of Federal Rule 403, which prohibits the introduction of otherwise-admissible evidence when “its probative value is substantially outweighed by

unfair prejudice” (Faigman et al., 2003, p. 3). As the authors viewed it, the potential worth of polygraphs as evidence was outweighed by the potential risk of skewing jury decisions.

Han (2016) explored the reasons for differing state standards on the polygraph. His numbers on usage were likely to still be current, and he indicated that 18 states allow conditional introduction of polygraph evidence, while the other 32 flatly forbid it. Every state that allows the usage of polygraph evidence requires that both parties stipulate specifically how it would be used and its relevance to the topic. Han (2016) theorized that there are a few reasons why states would require joint agreement on the topic. First, if both the prosecution and defense agree to the admission of a polygraph test, this is an implicit acceptance of the veracity of the evidence. By allowing it where it is frequently prohibited elsewhere, the parties are indicating sufficient trust in its validity to allow jurors and judges to interpret it for themselves. Second, if the parties agreed to its use, this could also be considered a waiver of their right to object based on evidentiary quality. The stipulation obviously did not change the accuracy level of the polygraph, but agreeing to allow it waived the parties’ right to object to its probative value (Han, 2016).

Some states tack on additional requirements in order for polygraphs to be admissible in court. Indiana Supreme Court rulings have held that, not only must the prosecution and defense agree to its inclusion, judges are mandated to consider the potential consequences of a wrong polygraph, because “the probable impact of the polygraph reference upon the verdict is of prime importance” (Han, 2016, p. 5). The side petitioning for the polygraph introduction must also demonstrate, via expert testimony or otherwise, a measure of reliability and acceptance of the polygraph and a clear

explanation of the theory and its implementation in the present case. Three of the 18 states which allow polygraph evidence restrict its use to corroboration, cases of impeachment, and probation revocation (Han, 2016, p. 5). The Washington Supreme Court held, based in part off a ruling from the Arizona Supreme Court, that polygraph evidence has an acceptable niche in terms of corroborating other forms of evidence or in impeachment of a defendant's testimony. If the defendant does not testify, 'standalone' polygraph evidence is prohibited (Han, 2016).

Part of the reason for skepticism around polygraphs in the courtroom was derived from the inconsistent methods of interpretation. Being found to be non-deceptive could be a strong influence on the justice system to drop its case against the individual. However, because of the differing policies of various agencies, the consequences, especially of inconclusive tests, could be wildly different. Honts (2013) explained that, in the FBI, an inconclusive result is treated precisely the same as a deceptive result, resulting almost inevitably in further interrogation and investigation.

Polygraph Law and Court Challenges

Polygraphs occupy a tenuous position in the legal system. Honts and Perry (1992) discussed the then-current laws in place regarding the usage of polygraphs both inside a courtroom and outside of it. Polygraphs have been in the past and still are components of some criminal investigations. Exams are occasionally used as tools by prosecutors and police agencies. Both parties may offer an accused individual an opportunity to take the polygraph; if clear, the charges would be dropped, but otherwise the case would proceed forward with the exam results likely to be referred to as evidence (Honts & Perry, 1992). Until the *Daubert* case in 1993, the measure for polygraph admissibility was the *Frye*

test, also known as the general acceptance test. If the methodology or technique used to obtain the evidence had obtained 'general acceptance' in the scientific community, it would be admissible in criminal proceedings (*Frye v. United States*, 1923). However, based on the ruling in *Daubert v. Merrell Dow Pharmaceuticals* (1993), the standards for evidentiary inclusion were changed significantly. The *Daubert* test required that the methodology used to gather the evidence be demonstrably scientifically valid and its relevance be properly explained to the matters being considered (*Daubert v. Merrell Dow Pharmaceuticals*, 1993). Even under the *Frye* test, polygraphs had encountered mixed results in terms of admissibility. Under the *Daubert* standards, only 18 states allow admission of polygraph evidence, and of these states, many require that the involved parties stipulate this beforehand, as well as allow the trial judge to be the ultimate decision-maker (Han, 2016).

Given that the *Daubert* standard requires polygraph evidence to be demonstrably valid and relevant, the lack of consistency in the literature is an issue for proponents of the polygraph such as Ben-Shakhar et al. (2002). Additionally, if the examiner needs to testify before the court, there are regulations that they must meet (Federal Rules of Evidence 702, 2018). A polygraph examiner is considered an expert witness in most cases. They did not see the crime committed, but their expertise in a matter relevant to the crime makes them capable of informing the court on the matter (Federal Rules of Evidence 702, 2018). As expert witnesses, in order for them to be allowed to testify, they must demonstrate to the court their credibility and ability to speak candidly on the topic at hand. The Federal Rules of Evidence (FRE) give the trial judge a great deal of discretion in determining whether the expert meets their standards for credibility (Honts

& Perry, 1992). Expert witness testimony must also meet the *Daubert* standard. In *General Elec. Co. v. Joiner* (1997), the court specified that results of the expert's analysis must logically lead to the opinion they give. Similarly, the case of *Claar v. Burlington* (1994) required that expert testimony account for the presence of obvious alternatives.

In cases that involve a polygraph, the lack of scientific support often requires the presence of an expert witness to convince the judge of its utility (Faigman, Fienberg, & Stern, 2003). It is important that the jury understands that the polygraph is *not* infallible, but emphasizing this potential error must be done without unfairly prejudicing the jury against the evidence, or the results may be dismissed (Federal Rules of Evidence 403, 2018).

However, even if the expert meets all the qualifications to the satisfaction of the judge, it does not necessarily mean that evidence they propose to explain would be permitted. Honts and Perry (1992) clarified that there are two types of evidence that polygraph results could potentially fall under: legal relevance and logical relevance. Legal relevance involves the question of whether the evidence tends to make more or less probable any facts of consequence in the matter. Logical relevance is slightly more abstract, and comparisons can be drawn to the *Frye* test. Logical relevance concerns whether the evidence being offered has a tendency to prove what it purports to prove (Honts & Perry, 1992). For example, if an attorney proposed some theoretical connection between consumption of asparagus and personality disposition to commit robbery, the judge would understandably have difficulty seeing the logical relevance of the defendant's dietary habits.

There are multiple reasons why polygraph evidence has fared differently under the newer standard than other types. While the *Frye* test requires general acceptance and *Daubert* does not, the shaky case precedent of polygraphs continues to be a problem. Judges who reject the polygraph in the states that do not universally prohibit it could point to a 2003 publication by the National Research Council, which concluded that improvements in polygraph accuracy have been slow and that the physiological measures employed by most polygraphs are inherently ambiguous (Pettit, 2007). Brain imaging techniques are significantly more accurate but also carry with them the problem of increased cost and the potential for the improper use previously mentioned. Another significant obstacle to polygraph introduction is judges' concerns that it "invades the province of the jury" (Pettit, 2007, p. 10). Under FRE 403, even relevant evidence may be excluded if its contribution toward establishing fact is mitigated by a substantial risk of prejudicing the jury (Federal Rules of Evidence 403, 2018).

Pettit (2007) also went into greater detail on the implications of the case of *U.S. v. Scheffer* (1998). Justice Thomas, writing for the majority, focused on three primary issues in upholding MRE 707's prohibition of polygraph evidence. Pettit (2007) identified the first of these issues as the government's interest in ensuring that evidence permitted at criminal trials is reliable, on which the Supreme Court ruled that "there is no consensus within the scientific community that the polygraph test is reliable" (Pettit, 2007, p. 11). The simultaneous presence of cautious criticism from Staunton and Hammond (2011) and strong support from Ben-Shakhar et al. (2002) on the veracity of the GKT is one example of what the Supreme Court may have been referencing. The second issue contributing to the affirmation of MRE 707 was a section of the majority opinion on

which Justice Thomas had only three other Justices join: Chief Justice Rehnquist, Justice Scalia, and Justice Souter (Pettit, 2007). Within this section, Pettit (2007) explained the Justices' interest in preserving traditional roles for criminal justice decision-making. In the opinion of these four Justices, "the jury *is* the lie detector" (Pettit, 2007, p. 12) and allowing the polygraph into courts poses a risk of changing this fundamental role of the jury. The third issue, again joined by Justices Rehnquist, Scalia, and Souter, was a government interest in "avoiding litigation over issues other than the guilt or innocence of the accused" (Pettit, 2007, p. 12). These four Justices believed that introducing polygraph evidence would result in additional lawsuits and legal questioning that do not relate to the crime at hand, which could potentially confuse the jurors as to whether they are determining guilt or proper examiner demeanor (Pettit, 2007).

The Supreme Court was far from unanimous in this case, as four justices on the majority did not join in the additional sections. Justice Stevens dissented with the entire decision. These five justices rejected the second and third arguments towards banning polygraphs. Justice Kennedy, joined by Justices O'Connor, Breyer, and Ginsburg, concurred that the scientific literature and community did not hold any consensus view on the polygraph, which was reason enough to uphold the prohibition in MRE 707 (Pettit, 2007). However, Kennedy and the other three justices specifically disagreed with the argument over ensuring traditional roles for criminal justice decision-making. They categorically rejected Thomas' argument that polygraph evidence threatened to usurp the jury's role as assessor and trier of fact (Pettit, 2007). Justice Stevens, the only fully dissenting justice, rejected the majority's claim that there was no consensus in the scientific community. Stevens did not argue that polygraph tests were perfectly accurate;

neither did he argue that they necessarily needed to be expanded in use. However, Pettit (2007) summarized Stevens' most vocal disagreement with the majority opinion as a flat rejection of the idea that polygraph evidence is seen as infallible by jurors. While not yet published at that time of the *Scheffer* decision, the findings of Myers, Latter, and Abdollahi-Arena (2006) lend credence to Justice Stevens' dissent.

Support for Polygraphs

Polygraphs are not without their proponents in the justice system and among published authors. Offe and Offe (2007) achieved a 90% rate of correct identification by using a CQT administered in a mock crime situation and opined that the justice system's view of polygraphs as inherently inaccurate is a criticism that cannot be sustained. They used a sample of both university students and law enforcement trainees, who were informed of the nature of the experiment. Offe and Offe (2007) also allowed participants to decide whether they wanted to participate in the survey as either guilty or innocent individuals. They incentivized participation by offering a 50 € reward for individuals who participated in the guilty sub-group and were deemed as innocent by the polygraph, as well as a 15 € reward for individuals who chose the innocent sub-group and were rightly exonerated by their polygraph. Ultimately, 35 people chose to participate as guilty and 30 chose innocent. While it (N = 65) was a reasonably-large sample, it was not large enough to make conclusions which can be widely generalized. However, as with all mock crime research, there exists the limitation of accurately simulating the severity of consequences that would be present in a real criminal proceeding. Failing a polygraph administered in a mock crime situation is only significant as a data point in the researcher's analysis. For a suspect under criminal investigation, failing a polygraph could determine whether they

are acquitted or convicted. The theoretical use for polygraphs in the justice system is virtually inarguable, as Ibek (2012) explained that it can indicate the presence of deception and guilt despite outward denial. Because few if any offenders genuinely wish to be caught and punished, they are unlikely to be forthcoming, and this is where an accurate polygraph exam could be an incredibly useful tool.

To improve polygraph accuracy and strengthen its theoretical background, Handler, Honts, Krapohl, Nelson, and Griffin (2009) had several suggestions to increase polygraph accuracy and generalizability. Handler et al. (2009) found a general trend among published writers that urged the expanded use of computers and data-driven hypotheses, very similar to the conclusions of Honts and Quick (1995) as well as Raskin and Kircher (2014). There is reason to be hopeful, as Handler et al. (2009) explained that the American Polygraph Association (APA) and American Society of Testing and Materials (ASTM) are working on forming a set of data-driven policies and guidelines for polygraph use everywhere.

Handler et al. (2009) designated several procedures for devising appropriate polygraph questions. Naturally, the answers to these questions should provide interpretable and useful information to the examiner. Questions should deal specifically with one alleged behavior or action. Increasing the length of time that a test subject has to remember back causes issues of telescoping, where individuals will mistakenly include events that were not within the timeframe specified (Handler et al., 2009). They should be easily answerable with just 'yes' or 'no', should not include any legal or criminal jargon, and should be devoid of references to motive or intent. As with every other aspect of the judicial system, the test should not presuppose the guilt or involvement of its

subject. Behaviors or actions need to have an operational definition that is understood clearly by both the examiner and subject. When presented in court, the results should also be explained intelligibly and their significance / probative value be understood by all involved parties (i.e. judge, prosecution, defense, and jurors) (Handler et al., 2009).

Despite other researchers looking doubtfully on the utility of polygraphs, Ben-Shakhar et al. (2002) believed that the GKT can be used successfully in criminal proceedings. The GKT is much less common, even for the already rare use of polygraphs in criminal proceedings, as the authors described that it is only used in Japan with any regularity. A brief review of the practice of the GKT is important here, as its differences from the CQT are what Ben-Shakhar et al. (2002) theorized make it more valid. Ben-Shakhar et al. (2002) believed one of the advantages of the GKT is that it focuses more on what an individual knows rather than trying to subjectively interpret meaning from emotions and conditioned responses. Their conclusion regarding the GKT polygraph exam was that it can detect neurological arousal / change even when no verbal response is given (Ben-Shakhar et al., 2002).

Ben-Shakhar et al. (2002) included two major requirements in order for GKT questioning and exams to be valid. For the orienting response to present itself, the questions need to pertain strictly to knowledge or details that a guilty individual would have and be likely to remember. For example, it would be very unlikely that a bank robber would notice and specifically recall the number of lights in the restroom of a bank they had robbed. Conversely, it would be much more probable that they would recall the combination or location of the keys to a cash drawer or safe. For a GKT question to be effective here, the examiner should focus on the crime-relevant issue objects. The authors

also made very clear that relevant information known to police should be kept stringently confidential so as to not ‘contaminate’ innocent subjects with guilty knowledge (Ben-Shakhar et al., 2002). This is an admitted weakness of the GKT, as in notorious cases, keeping all details from the public may be impossible. However, Ben-Shakhar et al. (2002) proposed a solution which seemed to remedy this at least in part. If the examiner was blind as to the relevant details, they would be much less able or likely to frame the questions in a biased way in order to get a conviction. The importance of preventing such contamination goes beyond the validity of the test. While very few cases, if any, have dispositions that are entirely determined by a polygraph result, the ‘fruit of the poisonous tree’ doctrine from an early Supreme Court ruling would require that a case which included a polygraph result be conducted in accord with all regulations and standards (*Silverthorne Lumber Company vs. United States*, 1920). The fruit of the poisonous tree doctrine states that, if initial evidence was gathered in an inappropriate or illegal matter, all evidence derived from it would also be considered illegal, even if the rest of the process was conducted in accordance with legal guidelines. A case which was based even in part off a polygraph could be destroyed by one incorrectly-administered exam (Ben-Shakhar et al., 2002).

Opposition to Polygraphs

Gallai (1999) argued that several factors regarding the then-current implementation and science of polygraphs render them inadmissible as evidence in a courtroom setting. The stipulations of federal law require replicable, non-prejudicial, and non-biasing circumstances for evidentiary submission. Under those iterations of federal

law, Gallai (1999) argued that judges are properly exercising their discretion to prohibit polygraphs as evidence in federal court.

Honts and Perry (1992) specifically analyzed one type of polygraph test: the control question test (CQT). CQT polygraphs ask two types of questions, those which pertain specifically to the crime or act under investigation, and a control question that deals with generally undesirable behaviors but not those under present investigation. At the time of publishing, Honts and Perry (1992) found a very large range in literature-determined accuracy for both CQT and GKT polygraphs, “ranging from chance to near perfection” (p. 5). Analyses / research that have been conducted on CQT and GKT polygraphs since that date have also been critical. An analysis by Honts and Raskin (1988) found rather strong support for the validity of the control question test: 85% of the innocent individuals were deemed so by the polygraph, and 92% of the guilty individuals were labeled such by their results. However, for the innocent group, the remaining 15% had polygraph results labeled ‘inconsistent’. For the guilty group, the remaining 8% were falsely exonerated by the polygraph. It must be mentioned that the sample sizes for both of these groups were extremely small (N = 13 for innocent, N = 12 for guilty), and as such, any conclusions drawn from the data should be approached with caution, as the small sample size makes it very difficult to generalize accurately to a larger population.

Honts and Raskin’s (1988) results were interesting when compared against other research on the CQT. In a later study, Raskin (1989) had more negative results: only 61% of innocents were seen as such by the polygraph, with 31% of this group having inconclusive results and 8% being falsely condemned. For the guilty control group, the numbers were not much better: 73% were properly determined to be guilty by the

polygraph and the remaining 27% were left with inconclusive results. The problem of sample size was slightly alleviated, as Raskin (1989) had an (N = 26) for the innocent group, and (N = 37) for the guilty test group. In a later study by Patrick and Iacono (1991), the negative trend was even worse: 30% of the innocent group was viewed as such by the test, with 46% inconclusive and 24% incorrectly condemned. The results for the guilty were surprisingly strong: 92% labelled correctly as guilty, and only 6% and 2% respectively for inconclusive and incorrect exonerations. Sample sizes were also improved in Patrick and Iacono's study (1991), though still not to the point where one would be comfortable theorizing widely-applicable conclusions from the data: (N = 37) for innocent, (N = 52) for guilty.

Honts and Perry (1992) alluded to one of the problems of polygraphs, that juries may have a much broader interpretation of the results and their significance than they truly represent. A polygraph result is ultimately the opinion of the examiner as to whether or not the individual being accused or testifying in court was truthful. The jury, as the trier of fact, determines whether the evidence presented to them by the polygraph is worth considering when the named individual is testifying (Honts & Perry, 1992). When considering the issue of logical relevance of polygraph results, the frequent criticism in literature and relatively large range in stated accuracy very likely contribute to the technique's uncommonness in the courtroom. Jurisdictions more influenced by the lower results for accuracy tend to prohibit it, while those that take greater stock in the higher rates err on the side of allowing it (Honts & Perry, 1992).

It is not solely criminal defendants and their attorneys who complain about the usage of polygraphs in the courtroom. Prosecutors often decry the use of polygraph

evidence in favor of a defendant, citing the hypothetical influence of a friendly polygrapher (Honts, 1997). Friendly polygrapher theory supposes that the examiner creates a testing environment where the accused individual is capable of answering any question to best suit their case, such as not asking substantive questions of an accused robber, but attempting to claim a positive, non-deceptive result as indicative of innocence. Because the procedures of polygraph exams vary by agency and jurisdiction, as well as by examiner, prosecutors claim that any individual can find an examiner who would be able to give them exactly the results they need for a case (Honts, 1997).

Despite this claim, Raskin and Kircher (2014) found little evidence to support this hypothesis and took issue with several of its presumptions. The authors disagreed that a so-called 'friendly polygrapher' would be able to entirely mitigate the stress and fear associated with polygraphs as no known jurisdiction allows differential reporting to courts (i.e., only reporting in the instances that suit one party). Individuals are advised that the results can be used even if it adversely affects their case before the pre-test interview takes place. Furthermore, even in these supposed friendly polygrapher environments, the subject of the test still has a great deal at stake. Results that indicate deception make it impossible for them to be used as a basis to dismiss the case. Individuals may also fear further increasing their legal costs and could quite easily be concerned about disrupting the relationship they have with their defense attorney (Raskin & Kircher, 2014). In order for an individual to 'beat' a CQT polygraph, they would have to show comparable rates of reaction to the control questions as they would to relevant questions, and the friendly polygrapher hypothesis fails to indicate how the purported

lack of concern would lead to the needed increase. All told, there is little support for the complaint, and it should be rejected as a reason to prohibit polygraphs (Honts, 1997).

Honts and Quick (1995) discussed the permissibility of polygraph results in terms of the Fifth Amendment right against self-incrimination. Even if the individual expressed a willingness to take the exam, it could still be prohibited due to not being direct 'physical evidence' in the same vein as a pair of stolen jeans or on-site sobriety tests (Honts & Quick, 1995). There is not a Supreme Court case that delineates which class of evidence polygraphs best fit. In the North Dakota case of *Wahpeton v. Skoog* (1980), the appellate court ultimately ruled that field sobriety tests are not protected by the Fifth Amendment, such that individuals may not invoke their right to silence or refusal without penalty. Because polygraphs do not deal with physical evidence, the North Dakota court continued in *Wahpeton v. Skoog* to hold that they could be prohibited under the Fifth Amendment without incurring penalty (Honts & Quick, 1995).

Similar situations arise in post-conviction or incarceration-based hearings. In the case of *Varnson v. Satran* (1985), marijuana was found in the possession of one inmate during a search, and the inmate was promptly punished and lost his 'good time' credit. As part of the decision-making process, the parole board required that the inmate take and pass a polygraph to remain eligible for parole. The examiner, based on the results and their interpretations thereof, believed the inmate to be deceptive in the second of two polygraphs. The inmate appealed the loss of parole and explained that, in his belief, the parole board's sole use of the polygraph as a factor in making the decision was impermissible (Honts & Quick, 1995). The appellate court ruled in favor of the parole board, finding that it was not unconstitutional for a polygraph exam to be made a

condition of parole. They also opined that the inmate's contention of inherent inaccuracy and imperfection of the polygraph was erroneous and found that the decision had not been made solely on the basis of the polygraph results. This is similar to the standards suggested earlier; polygraphs can be considered as evidence if the results are able to be methodologically triangulated (Handler et al., 2009).

Two main challenges arise where polygraph evidence is concerned. The first of the two calls into question the validity / competence of polygraph training. Honts and Perry (1992) found that polygraph examiners often lacked a background in psychology. Understanding interpretations of instrument-given statistics was also lacking, even in the Department of Defense Polygraph Institute, or DODPI (Honts & Perry, 1992). It would be unfair and inaccurate to imply that no such improvements have occurred since that time. Indeed, changes have taken place and federal polygraphs are now overseen by a larger agency known as the Institute for Credentialing Excellence (ICE). This agency provides, in addition to training and accreditation resources, a national standard for areas of work which require specialized training or education. ICE is responsible for development and improvement of polygraph training and has legal authority to conduct audits of agencies to ensure they comply with the standards for continuing education and scientific rigor (Institute for Credentialing Excellence, n.d.). A problem that is more likely to still exist is the issue of a lack of standards across agencies. Part of this is no doubt a product of the fragmented structure of law enforcement agencies in the United States. Establishing a uniform set of procedures would likely require a federal court order, due to the influence of inter-agency politics and the de-centralized nature of American policing. Because of the influence of American political federalism, each

police department can have its own standards and be tailored to the desires and attitudes of the area it serves (Katz, 1997).

The second issue concerning polygraph evidence is the possibility of countermeasures of various sorts. The most frequently-cited countermeasure is drugs. Theoretically, the alterations resulting from substance intake could ‘confuse’ the instrument and render the results inconclusive or inapplicable (Honts, 2014). However, even a cursory search on most academic databases discovered that there is a severe lack of information and research on drug countermeasures to polygraphs. Therefore, it was difficult to say with much certainty as to what effects drugs may have on polygraph efficacy. In terms of relatively obvious / basic indications of drug use (such as drowsiness or visible intoxication), Honts and Perry (1992) clarified that, not only do polygraph examiners ask beforehand whether an individual has consumed / used anything prior, the instruments are often capable of detecting what they are designed to measure even with a polygraph-beating substance present. Far more plausible as a problem is the presence of learned countermeasures.

Honts and Perry (1992) found that, if individuals were trained in the methodology and science of polygraphs, they were much more capable of beating the exam. Individuals were taught to recognize control and relevant questions, were also given suggestions of subtle physical / mental countermeasures (e.g. mentally reciting the alphabet, or physically biting one’s tongue), and they used these in the control question section of tests. Employment of these learned countermeasures skewed the ‘baseline’ response and therefore made it difficult for the examiner to detect deception. In one such experiment, all the individuals who received this psychological training were able to beat

the exam. Typical polygraph examiners do not give these more legitimate countermeasures adequate weight (Honts & Perry, 1992).

Countermeasures are not exclusively the tools of criminals. AntiPolygraph.org is a website that not only publishes articles discrediting and criticizing the polygraph, but also provides a 220-page book in digital form that explains various types of countermeasures in great detail (Maschke & Scalabrini, 2005). The first types of countermeasures that the book explained pertain to breathing. According to this book, polygraph examiners often have a 'baseline' of acceptance of anywhere from 15 – 30 breaths per minute, with about three seconds between each exhalation. To throw off an examiner, one way to manipulate breathing is to simply block it shortly (i.e., hold one's breath after exhaling for 4 – 5 seconds). The length of time that this block is present is crucial, as too long of a pause will likely be characterized as suspicious by the examiner (Maschke & Scalabrini, 2005). Another method of manipulation is decreasing the breathing rate for 5 – 15 seconds, ending before the next question is asked. For mental countermeasures, the authors suggested performing mental calculations or imagining oneself in extremely stressful or exciting situations (e.g., watching a volcano erupt, counting backwards from a random number in random intervals). Maschke and Scalabrini (2005) indicated that this should be done upon recognizing a control question or immediately after answering a control question. Physical countermeasures could be as simple as biting the tongue hard enough to cause pain, but not enough to bleed. This is marginally more difficult to do without being detected, but Maschke and Scalabrini (2005) specified that it too could have an effect on the accuracy of the polygraph. None

of the countermeasures listed by Maschke and Scalabrini (2005) required any specialized training or seemed impossible for a layperson to learn and employ.

A more recent study by Honts (2014) specifically looked at countermeasures to the polygraph and analyses of their effects. It is likely that counters to the polygraph have greatly changed since the 1990s, and as such, the 2014 data set made it easier to apply towards a modern law enforcement agency. It is important that data continue to be collected on polygraphs and deception, because as Honts (2014) pointed out, the conflict between deceivers and those who seek to detect this deception is an ongoing 'arms race'. Unfortunately, Honts (2014) indicated that much of the recent research on polygraphs had been done under the shroud of government classification and as such was not available for reference or discussion.

One of the concepts analyzed by Honts (2014) was the utility of spontaneous countermeasures. These are what they sound like: on-the-spot attempts by the individual being questioned to circumvent or skew the results to inconclusiveness. A spontaneous countermeasure is any behavior or practice that an individual may employ in the belief that it will skew or render inconclusive the results of the polygraph, such as clenching the jaw or fists, or deliberately trying to think of nothing. Because spontaneous countermeasures are not often devised with the type of exam or instrument in mind, they are often easily detectable by examiners. Honts (2014) concluded that these were largely ineffective, both because the theory behind such countermeasures was lacking as well as examiners recognizing and addressing the countermeasure immediately. In a sub-study referenced, Raskin and Kircher (2014) found that among the guilty, usage of spontaneous countermeasures did not significantly raise the percentage of falsely negative tests. The

problem was magnified for innocent individuals who used one or more countermeasures: usage significantly increased their scores in a negative direction (Honts, 2014).

Spontaneous countermeasures merit investigation, and the frequency of their usage that Honts (2014) detailed only emphasized the necessity to understand them. Previous research on this topic conducted by Honts, Raskin and Kircher (1987) analyzed the effectiveness of these spontaneous countermeasures. Guilty participants in these trials committed a mock crime and were administered a CQT about the mock crime. The total sample size in this experiment was 65, which is still relatively small and is cause for concern in regard to validity. Of the 65 subjects, 39 self-reported utilizing spontaneous measures to beat or distort the exam, a rate of 60% (Honts, Raskin, Kircher, and Hodes, 1988). The polygraph evaluators who administered the exams had a rate of 80% correct identification, 3% incorrectly identified, and 17% inconclusive. In the opinion of Honts et al. (1988), while spontaneous countermeasures were common, they were ultimately rather ineffective in obscuring CQT results.

A later study by Honts, Amato, and Gordon (2001) also analyzed the usage and viability of countermeasures. An advantage that this experiment had over the previous (Honts, Raskin, and Kircher, 1987) was that it also contained a control group of innocent individuals who did not commit the mock crime. Participants, similarly to the previous study, were administered a CQT pertaining to the crime, after which the examiner determined whether the individual had or had not been truthful. The sample size also improved, from 65 to 192, which was beneficial in terms of validity and reliability. Of this larger sample, half of the group was designated to be innocent while the other was guilty (Honts et al., 2001). The findings on spontaneous countermeasure use were similar

to the older study (Honts et al., 1987). Out of the 192 test subjects, 130 reported using some form of countermeasure in an attempt to obscure the results of the polygraph, a rate of 67.7% (Honts et al., 2001). Participants in the innocent test group reported relatively high usage of countermeasures, with 44 out of the 96 (equal to 45.8%) indicating that they had used one or more countermeasure. It was clear from these findings that spontaneous countermeasures did exist and that they were utilized by both the innocent and guilty. As expected, those within the guilty test group reported much higher rates of countermeasure usage, with 86 out of the 96 (equal to 89.6%) subjects (Honts et al., 2001). Among both groups, mental countermeasures were reported to be the most common tool to circumvent the polygraph. Three instructors from the Department of Defense Polygraph Institute (DODPI) were asked to evaluate the exams not only to determine truth or deception, but also to report on whether they detected usage of countermeasures. Ultimately, Honts et al. (2001) found that detection of countermeasure use was far less promising than detection of deception. Between the three evaluators, only 44 identifications of countermeasure use were made out of the entire sample. Honts et al. (2001) summarized the state of countermeasure detection critically, saying that “none of the DODPI evaluators detected countermeasure use at better than chance levels” (p. 7). Not only was there a low level of detection among the evaluators, inter-rater agreement on identifying countermeasure use was also very low; that is, the evaluators often disagreed on which subjects were employing countermeasures (Honts et al., 2001).

Ben-Shakhar et al. (2002) also addressed the incidence and utility of polygraph countermeasures on the GKT when compared against the CQT. The problem of learned countermeasures that Honts and Perry (1992) explained was the predominant issue that

Ben-Shakhar et al. (2002) discussed. Especially today, with the availability of websites and information on polygraphs, it is quite easy to access information which theoretically could teach an individual how to beat a polygraph, and this is corroborated by the frequent theme in research of the power of education (Honts & Perry, 1992). Honts and Perry (1992) directly showed the potential skewing power of education on polygraph test administration. The importance of questions and responses was equal for both examiner and examinee.

Honts (2014) found support for the power of misinformation on beating the polygraph. In this context, misinformation was the introduction of incorrect information occurring during the pre-test interview. For example, if a polygraph was administered in an investigation for a crime where the primary evidence was an eyewitness account, it was possible that the eyewitness may not remember the scene correctly (e.g., the getaway car was really red, but the witness reported it was gray). If this was the account the polygraph examiner was given and test questions were related to it, this would decrease the accuracy of the polygraph. In this hypothetical situation, even if the test subject did commit the crime, he / she could truthfully deny having seen the gray getaway car. Misinformation can also be done deliberately, where an examiner would develop test questions that emphasized a point the examiner knew was wrong. If the test subject's memory was correct, yet they were repeatedly being detected as deceptive in regards to this detail, it would be possible that they relented and conceded that the examiner's version of events was correct. Deliberate misinformation is less common, as it could be viewed as a form of entrapment against the individual. It is unsurprising, given criminology's understanding of the flaws in human memory (Walsh, 2012), that having

erroneous information introduced could skew accuracy levels, even if the individual directly committed or witnessed the crime. Solely introducing misinformation had a significant impact on the levels of polygraph accuracy. In a study by Amato-Henderson, Honts, and Plaud (1996), introduction of misinformation resulted in a majority of guilty respondents incorrectly being labelled as truthful by the polygraph. Innocent suspects who happened to possess crime-relevant information (frequently by witnessing it themselves) also were endangered by the introduction of misinformation (Honts, 2014).

According to Rovner (1986), introduction of pre-exam information influenced polygraph accuracy rates. As other research has indicated, individuals who were well-versed in the sort of exam they were about to be subjected to often were better able to be deceptive without being detected. While an examiner frequently asks what the subject has heard or knows about polygraphs, they do this primarily to counter arguments or allay the fears of the subject. It would be unlikely that trained examiners would disclose too much of the formatting of the test because they know of the risks, but Rovner (1986) found that this pre-test disclosure and conversation may itself be a risk factor in terms of accuracy. While justice system transparency is not a problem, the wider availability of information and required disclosure from one party to another does bring with it greater public understanding of basic concepts of polygraph theory and testing formats.

In scientific experimentation, it is useful to separate the experimental group into two sub-groups. In Rovner's (1986) study, the control group was not given any relevant materials before the administration of the test, but one of the experimental subgroups was given a loose-leaf notebook with detailed information regarding the theory of the test, the data / patterns that an examiner may use to make their decision, and a list of

countermeasures that could potentially skew the responses. The second subgroup was given access to this notebook of information, but was also told what model of polygraph instrument was being used and what it detected, as well as given two practice tests before the 'guilt' test. The question order remained the same on both practice tests. After these practice runs, the second subgroup also was allowed to examine the results and discuss them with the polygrapher (Rovner, 1986). The author had the results of the official exam scored by a third examiner who was not present during any of the tests.

Rovner (1986) discovered that solely being presented with information on the theory of the test had no significant effect on accuracy levels. Rates of accuracy for both the control and information groups were approximately 91.5% correctly identified, 8.3% incorrectly identified, and only 0.2% gave inconclusive results. However, for the group who was given information in addition to two practice exams, the rates were significantly lower: 70.8% correct identifications and 29.2% incorrectly identified. It is true that 70% is still a rather large majority, but any concerns present with even an 8.3% error rate should significantly increase with an error rate almost four times higher. Rovner (1986) found a high level of interrater reliability for his methodology via comparing the decisions made by the examiners and the independent evaluator. Both had very high levels of definite decision-making, 93% for the administering examiner and 90% for the evaluator. On the subjects analyzed by both an examiner and the independent evaluator, the rate of agreement rose to 95%. Numerical result comparisons between the examiners and evaluators had an extremely strong correlation: ($r = 0.97$) (Rovner, 1986).

For future researchers, Rovner (1986) had a few suggestions and conceded some limitations to his study. One factor that remains true today and applies to virtually all

crime research is the necessity of creating a realistic scenario for the test subject.

Understandably, there cannot be realistically severe consequences, but a bit of deception appears to be necessary in shaping a realistic environment. Rovner (1986) incentivized taking the test by offering a small cash compensation for participation as well as an additional reward for those who successfully managed to beat the polygraph.

Amato-Henderson, Honts and Plaud (1996) researched the impact of misinformation on guilty knowledge test (GKT) polygraphs. The total sample size was between the older study (Honts et al., 1988) and the more modern Honts (2014) study, at 96 participants. The methodology was slightly different than the other studies. Instead of committing a mock crime, test subjects watched a videotaped crime in order to induce guilt (Amato-Henderson, Honts, & Plaud, 1996). After a week, test subjects were given deliberate misinformation regarding certain facts of the crime, administered a GKT polygraph regarding three misinformed facts, three un-altered facts, and after completion, all respondents took a 20-question test regarding the crime. Misinformation was determined to be successful if it resulted in a higher likelihood of being designated as innocent (Amato-Henderson, Honts, & Plaud, 1996). Logically, it stands to reason that introducing erroneous information could negatively impact the veracity of a polygraph exam, and this was the result that Amato-Henderson et al. (1996) discovered. On the GKT polygraph, 52 of the 92 test subjects (56.5%) were incorrectly identified as being truthful about their innocence, while only 40 of the 92 (43.5%) were still correctly identified as guilty. From this data, it was readily apparent that misinformation impairs the accuracy of a GKT polygraph exam (Amato-Henderson et al., 1996).

The length of time that passes between the crime and the polygraph exam is also a risk factor against accurate results. Every recall of a memory places it at risk for alteration (Honts, 2014). Over time, various forms of misinformation could sufficiently alter a memory so that it is completely devoid of fact, though the individual does not believe it to be. While no pharmaceutical company is likely to produce a drug specifically tailored to help people circumvent the polygraph, Honts (2014) cautioned that the side effects of certain drugs help degrade the memory quality of certain events. If an individual was facing a polygraph and knew they had relevant knowledge, all they would need to do to bolster their chances of circumventing the polygraph is to recall as much as they can regarding the event, take the prescribed drug, and wait. The memory is all but guaranteed to degrade in quality or possibly even be erased. However, such substances should not be considered solely criminal aids, as Honts (2014) spoke of their use in treating post-traumatic stress disorder and drug addiction. Further research is needed to determine the levels of influence these pharmacological drugs have and how best to screen for them.

When it comes to detecting the use of countermeasures, the numbers are disturbingly low. Part of the problem is that, as soon as one particular countermeasure is detected and accounted for, another one may take its place (Honts, 2014). More so, it is very unlikely that accurate counter-countermeasures can be developed for mental countermeasures, since the medical field does not fully understand or have a way to fully pictorialize the process of recall. Relying solely on examiners to identify usage is demonstrably inadequate, as evidenced by Honts et al. (2001). Worse so, Honts (2014)

illustrated that as much as 48% of the individuals accused of using countermeasures were actually innocent.

While it is conceded that polygraph instruments can detect what they are designed to, Honts (2014) also indicated examiners could be inadequate when it comes to detecting deceptive behaviors. According to Honts (2014), some of what examiners may interpret as overt signs of deception (e.g., shuffling feet, shifting eyes) is actually natural behaviors displayed by innocent people. For anyone who has sat through a polygraph, even in a non-criminal context, the procedure is often very stressful. An individual who feels nervous in a polygraph situation may fidget or display other physical signs of nervousness, and if the instrument identified these movements as deceptive, it is possible that an innocent individual could be wrongly accused of committing a crime or lying to the polygraph examiner. Unless the U.S. government adopts a more “enlightened attitude” towards polygraphs and polygraph research, Honts (2014) saw little likelihood that scientists would be able to make much progress.

Honts and Quick (1995) explored the disadvantages of the most common type of polygraph exam (CQT). The level of disagreement over the merits and flaws of the CQT was described as a “polemic controversy” (Honts & Quick, 1995, p. 6). The CQT exam asks test subjects two types of questions: those relevant to the matter under investigation and those that are not. The test suggests that innocent individuals, because they do not have any knowledge of the subject of investigation, will have greater physiological responses to the control questions (Honts & Quick, 1995). Detractors of the CQT, at the core, accuse the test premise of being naïve. Critics claim that no examiner can formulate

control questions with as much psychological impact as relevant questions to serious crimes (Honts & Quick, 1995).

Much more interesting were Honts and Quick's (1995) conclusions regarding the issues of polygraph research. While laboratory research is often desirable, it can be difficult, especially where crime is concerned, to truly 'replicate' in a sterilized setting. Because of the lack of standards in polygraph examinations, the way that the researchers conduct the exam may be different than the way the local police agency conducts theirs. Additionally, it would be difficult if not impossible to replicate the seriousness of the consequences that are present in a real-world polygraph. Individuals in a lab experiment would be unlikely, even if specifically instructed, to treat the exam as if they were in a real-world scenario, which can influence the validity of the results. Furthermore, there is difficulty in ascertaining a baseline from which to compare responses to, especially when dealing with relatively abstract concepts such as truth (Honts & Quick, 1995). There are issues with confessions and polygraphs, and Honts and Quick (1995) explained that, in some cases, police can be more concerned with obtaining a confession rather than conducting a proper polygraph examination, which could diminish the sense of objectivity that accreditation agencies often have as a requirement.

Raskin and Kircher (2014) researched the issue of human error intrinsic in polygraphs. While the subjectivity in examiner interpretations of results was present already, the authors raised a valid question about the mental and emotional state of the subject. Given the differences in personality, it does not strain the imagination to think that a psychopath might have less of an issue lying. One proposed solution to this is allowing computers to score and interpret the majority of polygraph data, as well as to

gather and compare trends in multiple exams (Raskin & Kircher, 2014). This appears to be a very strong suggestion to increase the validity of results. Statistical aggregation allows researchers to draw inferences about trends that are not observable from one case or a small number. Raskin and Kircher (2014) found that computer algorithms have accuracy that is at worst comparable to skilled interpreters and at best better than these interpreters. Unsurprisingly, the one class of mental disorders that has an uncontrolled influence on polygraph veracity is full-blown psychosis. While these individuals may be better able to circumvent the polygraph, the other symptoms of psychosis (i.e., hallucinations, delusions, and catatonia) often make them identifiable during the pre-test interview. Raskin and Kircher (2014) found that, despite the face validity of psychopaths being able to lie on an exam, the actual results did not show much strength for this connection.

Handler et al. (2009) indicated that the lack of consistency in polygraph literature and methodology extends to the interpretation and conclusion-forming of results as well. The most common way this occurs is termed global analysis (Handler et al., 2009), where the examiner directly inspects the test results for physiological responses or trends in responses to a particular question or two. The limitations of this are obvious: it introduces examiner subjectivity and is almost entirely dependent on the examiner's ability to recognize visual patterns from the available data. Increasingly frequent is the utilization of numerical scoring (Handler et al., 2009, p. 7), in which examiners run basic statistics on the data to compare magnitude and assign objectively-defined scores, with cutoffs at certain levels specifically delineated as truthful, deceptive, or otherwise. In a criminal

investigation, an inconclusive result on a polygraph exam often results in a re-test, where the rates of inconclusive results tend to drop (Handler et al., 2009).

In looking at the 2011 American Polygraph Association (APA) meta-analysis data, Honts (2013) found that rates of inconclusive results were actually higher for innocent individuals compared to the guilty. That is, individuals who did not commit the crime had a higher probability of an inconclusive score than the person or persons who did commit the crime. The APA's reported rates for the most accurate and stringently-administered forensic polygraph exams were rather subpar: a false positive rate of 13.8%, false negative rate of 3.4%, innocent yet inconclusive rate of 25.5%, and guilty yet inconclusive rates of 3.4%. This must be considered with the caveat that all inconclusive results were treated as indicative of deception. Honts (2013) included this in his analysis, and while he found that removing all such cases did not have an influence on the confidence of a non-deceptive result, it decreased the likelihood of obtaining a true deceptive result to 71.3%. In his own words, treating all inconclusive results identically to guilty results ultimately "has no effect on the confidence in a truthful outcome" (Honts, 2013, p. 2). The additional interrogation of innocent individuals whose results were inconclusive had no statistical impact on how confident the ultimate decision may be.

With the push for judges and all participants in the justice system to make evidenced-based decisions, this is a rate that merits concern (What is EBDM?, 2017). Roughly 30% of the time, the polygraph may be wrongfully impugning an individual's character or recalled actions. While there is no specific number that determines what is permissible under the *Daubert* test, it seems doubtful that this would meet that criteria or

be considered something that could be trusted beyond a reasonable doubt as is the rule for criminal court decision-making (Honts, 2013).

While polygraph result interpretation is not strictly tied to court permissibility, it is also worth considering in terms of resource allocation. Because every criminal justice agency (even at the federal level) must contend with a limited quantity of dollars and employees, it is imperative that the practices employed by these agencies be effective, efficient, and financially viable. Honts (2013) found that the opposite was true for the inconclusive-as-deceptive approach to interrogation and criminal polygraphing. The number of innocent individuals who were needlessly interrogated due to an inconclusive result increased by an astonishing 190%, while the rates of guilty individuals who were brought in for further interrogation increased by only 15% (Honts, 2013). English jurist William Blackstone (1778) wrote in his magnum opus that, “the law holds it better that ten guilty persons escape, than that one innocent party should suffer” (Blackstone, 1778, p. 713), a goal that seems to be in direct conflict with this practice.

The polygraph is obviously in a tenuous position in terms of permissibility in the courtroom, since the standards and regulations vary across the states and levels of government. In the armed forces, the Military Rule of Evidence 707 forbids the usage of polygraphs in all court-martials, on the basis of *Frye* test failure and inaccuracy (Military Rules of Evidence 707, 2015). Upon initial draft in 1991, then-President George H.W. Bush interpreted the language as a “per se ban on all polygraph evidence in courts-martial”, and this interpretation was de facto accepted as law for seven years (Military Rules of Evidence 707, 2015). This issue was litigated to the U.S. Supreme Court in the

case of *U.S. v. Scheffer* (1998), where the Court held that such a ban on evidence did not violate the Sixth Amendment right of the accused to present a defense.

There exists one major caveat to this ruling that must be understood. In *U.S. v. Scheffer* (1998), the justices specified that the previous hard-line prohibition on polygraph was not necessary; judges could still make a determination on a case-by-case basis per *Daubert*. Currently, the U.S. military still abides by the Military Rules of Evidence, including 707, which technically bars the inclusion of polygraph evidence in proceedings (Military Rules of Evidence 707, 2015). Because of the inconsistency in the literature, it is doubtful that many judges in courts-martial are likely to make any changes or rulings that stray from 707 regulations. While there are, of course, differences between military courts and civilian criminal matters, this is certainly a relevant case law to consider. The polygraph is a singular part of the overall criminal justice system process, though its results can be felt in all three stages (i.e., police custody, court proceedings, and corrections). The method and environment of the police-administered polygraph may impact its permissibility and relative credence among the jurors, whose decision can then result in further litigation after trial dispensation. Thus, it is important that there be a basic understanding of the public's view of the justice system.

Perceptions of the Justice System

Hurwitz and Peffley (2005) analyzed the different perceptions of procedural and structural fairness of the justice system. American public opinion of the police has always been divided. In Hurwitz and Peffley's (2005) own words, "whites believe the criminal justice system is fundamentally fair, and most African-Americans do not" (Hurwitz & Peffley, 2005, p. 3). It is imperative to understand why such differences exist and what

influence they can have on the efficacy of the justice system. This concept is not unique to criminology and can be just as easily applied to the economic system in the U.S. as well: whites see mostly fairness and equal-opportunity, while African-Americans see neither (Hurwitz & Peffley, 2005). While understanding these viewpoints and what events develop them is important, Hurwitz and Peffley (2005) indicated that most research on the topic did not branch between both sides: i.e., data were collected from mostly whites or mostly blacks in single studies, with relatively few achieving equal interracial comparison. These comparative studies are useful as references, because they allow members of the majority and minority groups to see what contributes to these different opinions. It is the lack of understanding and inability to see the others' perspective that makes criminal justice policy and events as controversial as they sometimes are.

Hurwitz and Peffley (2005) drew data from the 2005 National Race and Crime Survey, which was a nation-wide random-digit telephone survey. The total sample size was large enough to mitigate concerns over representativeness or generalizability (N = 1182). Not surprisingly, the data clearly supported the initial hypothesis on greater skepticism amongst African-Americans compared to whites. Over 67% of the African-Americans surveyed believed that the justice system does not treat people fairly and equally, where only 35.2% of whites shared this cynicism (Hurwitz & Peffley, 2005). These broad opinions of the justice system influenced the way that respondents perceived following situations in the survey. When asked to judge a hypothetical situation of police brutality, white respondents tended to judge slightly less punitively when the victim was white as opposed to black. Similarly, in a stop-and-frisk situation, whites were marginally

more likely to side with the police if an arrest was made. Blacks who believed that the system was fundamentally unfair were more likely to view all such hypotheticals skeptically, siding with the victims more often than the police. Perhaps most surprising of all, Hurwitz and Peffley (2005) found that whites have a rather naïve view of racial impact in the justice system. It is readily apparent that races are treated differently in the justice system, but whites in Hurwitz and Peffley's (2005) sample seemed unaware of this phenomenon.

Skewed perceptions are not restricted to ideas on racial discrimination. For example, many proponents of the death penalty believe that the sentence is faster and that it is not worth taxpayers' dollars to keep inmates alive for years on end, but this is demonstrably false because it does not take into account the cost (both time and money) of guaranteed appeals all death row prisoners are entitled to. Marcus (2007) indicated that, while the U.S. still has a relatively high number of individuals on death row, these numbers included those "whose appeals and legal petitions have been pending for many years" (Marcus, 2007, p. 10). Until these appeals are heard by a court, the state is prohibited from executing the inmate, even if the inmate refuses to file any appeals themselves. Death row prisoners in the U.S. can expect to remain incarcerated for several years before a major change in their case occurs, and all this time they incur the same costs of a life-without-parole prisoner (Marcus, 2007).

The public also has an incorrect idea of the purpose and environment of drug and specialty courts, believing that they are easy and soft routes for offenders to get out of the justice system quickly. In reality, the opposite is true: specialty courts often require a

vetting process for a candidate that analyzes their potential for reformation through alternative treatment (Goodfellow & Kilgore, 2013).

Unfortunately, individual perceptions of the justice system are out of touch. Kozinski (2015) opined that, while the public pretends otherwise, “much of what we do in the law is guesswork” (Kozinski, 2015, p. 1). One of the ways to counter these incorrect assumptions is through education. While no one can be compelled to educate themselves on criminal justice, those who select it as a field of study can be researched, and this was the relationship analyzed by Tsoudis (2000). She sought to determine the strength of the connection between choosing a criminal justice major and overall view of the justice system. To accomplish this, Tsoudis (2000) surveyed 200 students at a Midwestern college. In this sample, 99 respondents were criminal justice majors and 101 were non-CJ majors. The survey was administered during class and included questions on attitudes pertaining to crime, beliefs relating to crime, corrections philosophy, and the overall role of the justice system and its actors. To emphasize the importance of this study, Tsoudis (2000) highlighted that, while criminal justice majors are often the administrators, staff, and workers in the system, the larger public still has a fairly significant impact on criminal justice. Because legislators, at least theoretically, represent their constituents, they are likely to be sensitive to the desires of those voting individuals. Tsoudis (2000) gave the example of the death penalty, explaining that, despite there being a lack of empirical support for its supposed deterrent effect on crime, the public still looks upon the practice favorably.

Ultimately, Tsoudis (2000) discovered that those who were criminal justice majors held four significantly different beliefs regarding the justice system. First, they

were less supportive of harsh punishment and lengthy sentences. Second, criminal justice majors were more accepting of protection of criminals' rights before and after the trial process. Third, they believed strongly that criminal punishments should not vary by race. Fourth, criminal justice majors were more supportive of a separate juvenile justice system and different philosophical approach for juveniles than were other students (Tsoudis, 2000).

Perceptions of Crime and Sentencing

Roberts and Doob (1990) also analyzed the impact of media on public perceptions of crime. Because murder, rape and other extreme crimes often saturate the news cycle, the public is only ever exposed to the worst of crime and therefore has a much more severe picture of crime than really occurs. This, when brought up to legislators, results in the ultra-punitive set of penalties for many relatively minor crimes that are still common in the U.S. (Roberts & Doob, 1990). Roberts and Doob (1990) found that, while murder accounted for only 1% of all crimes, it was the topic for 25% of all crime stories. Among news media in the U.S., roughly one in five news stories covered violent crime, when only 6% of total crime was actually violent (Roberts & Doob, 1990). This was clearly out of touch with reality, yet it was the version of events that ~95% of people reported as their first source of information on crime.

The primary topic of study for Costelloe, Chiricos and Gertz (2007) was determining the extent to which this discontent and dissatisfaction was statistically tied to more aggressive policies on crime. A survey was conducted among 2,250 randomly-selected Florida residents, who were interviewed at various points between October and December 1997. The sample demographic closely mirrored actual state demographics in

every respect except for mean age, which was approximately four years older than the state average of 38 years of age. To measure opinion harshness, respondents were asked to rate on a 1 – 10 scale (1 being little to none, 10 being almost complete) their support for various punitive crime policies, including support for the death penalty, chemical castration for sex offenders, and limiting death penalty appeals. In measuring fear of crime, respondents stated whether they were fearful of being the victims of certain violent crimes (Costelloe et al., 2007).

Surprisingly, Costelloe et al. (2007) found that fear of crime was a stronger predictor of punitive opinion than actual criminal victimization among both males and females. Marriage was also found to be a significant predictor of punitive attitudes, especially for women. The data did not support the authors' hypothesis on economic insecurity and dissatisfaction among any age or sex group. Economic insecurity did have an effect on the attitudes of white respondents; when they expected their economic situation to worsen in the coming year, respondents held significantly more punitive opinions of the criminal justice system (Costelloe et al., 2007). The impact of education also varied between ethnicities. For black and Hispanic respondents, increased education led to less punitive views but harsher opinions among whites. In conclusion, Costelloe et al. (2007) agreed with the concept of the "angry white male" when it related to economic dissatisfaction leading to more draconian crime policy (p. 18).

The continuous exposure to heinous crimes from the media leads to the development of increased fear of crime. A consequence of increased public fear can be observed in the trend towards punitive sanctions. Costelloe, Chiricos, and Gertz (2007) examined other factors that contributed to the development of harsh public opinion. The

gravity of the situation was apparent, as Costelloe et al. (2007) illustrated that, between 1986 and 2006, the total population of incarcerated individuals rose by 77%. There were multiple legal practices that contributed to this enormous surge in jail and prison population. Costelloe et al. (2007) highlighted a few examples of such policies, including the increased frequency of juvenile waiver to adult courts, the expansion of “three strikes” and habitual offender laws, and extremely harsh sentences for even non-violent drug offenses. Despite these facts, 65% of Americans in 2006 felt that the court system was still too lenient on offenders (Costelloe et al., 2007).

Exposure and response to a violent crime-saturated media are only part of what could contribute to the draconian public perspective. Costelloe et al. (2007) hypothesized that individuals have become increasingly dissatisfied with the state’s inability to provide physical or economic security, and that this discontent finds an outlet in implementing harsh policies against a “ready-made, deeply unpopular target population”, that is, criminals (p. 2). Due to increasing deindustrialization, American companies face immense economic pressure to outsource jobs to locations where they are more profitable. The market for unskilled labor in the U.S. has dropped considerably since the 1980s and continues to recede today. Wages, benefits, and pensions have also shrunk in this time, and to no surprise, the workforce is largely displeased with these changes (Costelloe et al., 2007). In short, due to the loss of economic mobility, and a lack of other outlets for such displeasure, Costelloe et al. (2007) believed that the public used harsh criminal punishments as a proxy for success to compensate for their frustration in being unable to move up the social ladder.

Nevertheless, the findings of Costelloe et al. (2007) did not go entirely unchallenged. McCorkle (1993) emphasized that the public is not universally supportive of punishing certain types of offenders, most frequently juveniles and to a lesser extent the poor and minorities. The public is cognizant that income has a significant impact on an individual's progress through the justice system, and lack of income can unfairly prejudice or disadvantage a poor defendant. The existence of a separate juvenile justice system is evidence that society does not see all offenders equally (McCorkle, 1993). Because of their younger age and perceived malleability, juveniles enjoy a number of protections that adults do not receive, among them the closure of nearly all juvenile court proceedings to the public. Judicial records for adults are often publicly accessible and can be looked up on case repositories or at courthouses, where some states close juvenile records, only allowing them to be opened with a court order (Sha Sha & Fine, 2014).

McCorkle (1993) did agree that, overall, the American public has a strong desire for punishment of criminal offenders, but tempered this statement with the addition that many "appear to recognize the limits of a purely punitive response to the crime problem" (McCorkle, 1993, p. 2). To obtain a better picture of public sentiment on the criminal justice system, McCorkle (1993) conducted a telephone survey of 967 adults in the Las Vegas area. Response rates to this survey were low, as only approximately 41% of the total sample actually answered the survey questions. Despite the low response rate, the median characteristics of the sample (i.e., majority white, nearly equal distribution by sex, various levels of education and income) were close to the demographics of the area and similar to those used in previous public opinion research. Respondents were presented with an example crime situation and then asked for their opinions on various

justice system practices in use and whether they believed changes needed to be made in how such cases were handled (McCorkle, 1993).

While the overall findings were still rather punitive, the average opinion was nuanced in a way that is not adequately represented by the public official who promotes a straightforward “get tough on crime” policy. For less serious crimes such as drug possession, McCorkle (1993) discovered that his respondents were fairly supportive of rehabilitative policies and programs for this offender and did not believe incarceration was the only acceptable response. Even this support was moderated somewhat, though, as McCorkle (1993) simultaneously found that his respondents preferred rehabilitative efforts and programs to take place in correctional settings. Even in McCorkle’s (1993) time, prisons and other facilities were often plagued with their own issues of crime, which forced staff and educators to direct most of their efforts first towards maintaining control.

Rossi, Berk, and Campbell (1997) observed similar results to that of McCorkle (1993). The only area where public opinion appeared to differ sharply from practiced law was in the different punishments meted out to drug offenders based on the drug they were abusing. Rossi et al. (1997) used a national survey of American adults and compared their responses to certain parts of federal sentencing guidelines. Each participant was asked to judge approximately 40 vignettes covering a variety of crimes ranging from “felon improperly owning a firearm” to “kidnapping, with victim being killed” (p. 9). Perhaps promisingly, respondents in this study displayed no significant variation in sentences based on extra-judicial factors such as defendant’s gender or employment history. Even prior record, one of the strongest actual predictors of future offending, was noted to have

only a moderate impact on sentence rates. Average sentence length across all crimes was approximately seven years. Rossi et al. (1997) learned that the longest sentences tended to be given for street crimes where one or more victims were killed. However, there was almost no support whatsoever for long sentences against drug offenders. This increased tolerance applied across all drugs, which was particularly unusual when compared against the federal guidelines.

Perceptions of Courts

Sentence length is not the only aspect of court function that citizens analyze, as they previously have raised concerns over the courts' treatment of victims. The importance of having a positive public view of the courts is paramount, as there is a dependency between the two for optimal function. The public needs the courts to "punish offenders so that others are not victimized" (Kaukinen & Colavecchia, 1999, p. 3). The courts also require willing citizen participants to serve as faithful witnesses and objective jurors (Kaukinen & Colavecchia, 1999).

Wood (2009) explained that the importance of citizen participation and opinion goes beyond their functions within the system. To be clear, witness cooperation and juror work are some of the most important functions that citizens can serve in the justice system, because if they did not do so, it would be highly likely that the system would be unable to function. Public opinion and satisfaction are equally important to the government, because ensuring that citizens feel their voices are heard and responded to are primary ways that governments obtain legitimacy. With this in mind, a degree of back-and-forth communication between the state and its population is inevitable and desirable. Wood (2009) indicated that police departments must not only tailor their

policies and responses relative to the level of committed crime, but also must take into consideration the level of fear of crime.

Because of the demographic variety in the United States, it is not always accurate to identify one specific opinion as being *the* public opinion. Members of the public have attacked the courts because they believed that sentences were too lenient and did not sufficiently punish offenders (Kaukinen & Colavecchia, 1999). Blacks, in particular, expressed greater concern regarding the courts' fairness to offenders and victims than did members of other racial groups (Kaukinen & Colavecchia, 1999). Once again, income also had an impact on perceptions, where citizens whose household income was near the poverty line expressed similar skepticism regarding the equality of the court process, believing that the results of a case could be strongly influenced by income (Reiman & Leighton, 2013). Reiman and Leighton (2013) wrote extensively on the problems of economic inequality and its effects on the justice system, titling the book, *The Rich Get Richer, and the Poor Get Prison*. One of the most noticeable disadvantages of low income in the justice system that Reiman and Leighton (2013) noted was the extreme difficulty of making bond before trial. The potential for further problems from failure to make bond was clear: difficulty in retaining employment, disruption of family affairs and security, and restricted ability to meet with attorneys.

In contrast, polygraphs appear less prone to the unbalancing influence of income. Wealthy individuals do not possess inherent abilities to deceive or lie compared to poor individuals. However, as Honts (2014) and others stated, information and knowledge of polygraph techniques resulted in greater instrument inaccuracy. An individual's access to such information and resources would be susceptible to the biasing effects of income: i.e.,

individuals unable to make bond would be less likely to have time to specifically analyze polygraph exam formats.

Citizen Opinion on Polygraphs in Comparison

Myers et al. (2006) researched one of the foundational questions studied in this thesis: the comparative weight that jurors gave to polygraph evidence, compared to other types of evidence. The primary hypothesis was that jurors tend to consider polygraph results a very significant indication of guilt. They conducted a survey among a sample of 411 randomly-selected individuals in California. Each participant was given a mock trial vignette on which they decided guilt or innocence; assignment to each type of polygraph evidence was done randomly. A post-test questionnaire was administered to each respondent, involving guilt or innocence, as well as a probability of commission (Myers et al., 2006).

The research was much more in-depth than merely inquiring in regards to guilt and probability. Respondents first rendered a verdict based on the information from the trial vignette. They then graded their confidence in their verdict on a percentage scale from 0 – 100%. All of Myers et al. (2006) survey participants answered a question on the accuracy and usage of CQT-based polygraphs. For the members of the sample whose trial vignettes included polygraph evidence, they were asked how they viewed polygraph accuracy in general, graded on the same 0 – 100% scale. The group with polygraph evidence was asked on a five-point scale to grade its influence on their decision, which ranged from zero to four. All respondents were asked if they believed that polygraph examiners should be allowed in court to testify on clarifying the interpretations of the evidence; that is, whether defendants were truthful or deceitful as detected by the

instrument. Respondents answered where they believed the polygraph should fit in a courtroom setting. These responses were “sufficiently reliable to be the sole determinant”, “useful diagnostic tool when considered with other available information”, “of questionable usefulness, entitled to little weight against other information”, and “of no usefulness” (p. 8). To assess relative weight, Myers et al. (2006) asked respondents to rank polygraph evidence compared with four other types: fingerprints, DNA, eyewitness, and handwriting. Finally, respondents were asked to complete questions on non-criminal information, primarily regarding their demographics (Myers et al., 2006).

In testing the hypothesis that jurors defer heavily to polygraphs, the percentage of convictions was first analyzed (Myers et al., 2006). Out of the 411 participants in the study, those who received negative results (indicating the accused had taken and failed an exam), 71.9% voted to convict. For those whose vignettes indicated a positive result (accused had taken and passed a polygraph), 59.1% voted to convict. Of the group who had no polygraph evidence, 70% voted to convict. It is apparent that there was a difference between negative and positive results, but it was not immediately clear as to whether Myers et al.’s (2006) hypothesis was supported. A chi-square test of independence was conducted to determine whether the presentation of polygraph evidence had a statistically significant impact on verdict. Myers et al. (2006) stated in their analysis that these tests did not show statistical significance. The mere presentation of polygraph evidence did not significantly impact the decision to convict or not to convict. To corroborate this finding, the authors also conducted a one-way ANOVA on the verdict to determine if accessibility of evidence had an impact. The ANOVA showed that access to polygraph evidence was not statistically significant in predicting verdicts. It

was clear from the analyses of Myers et al. (2006) that accessibility to polygraph evidence by itself did not have a significant effect on verdict decisions. The issue in previous studies on polygraphs of sample size was also less problematic here, with (N = 411), which supported its internal validity. The use of a random public sample in Myers et al. (2006) increased its external validity.

In regards to the question measuring confidence in the polygraph, every respondent was informed of the polygraph having an estimated accuracy of 85%, which is relatively close to the 80% rate of correct identification put forth by Honts et al. (1988). Even with this information, both the failed-polygraph and passed-polygraph groups reported lower degrees of confidence in polygraph evidence: 68% and 62% respectively (Myers et al., 2006). The difference between these two ratings was not found to be statistically significant. When these participants were asked how the evidence impacted their decision on the five-point scale, there was not a statistically significant difference between the failed-polygraph and passed-polygraph groups. Myers et al. (2006) also tested whether education and age had an impact on perceived polygraph credence. For both variables, there was no significant difference between passed-polygraph and failed-polygraph groups. Education did not lead to higher skepticism towards the polygraph, and neither did older participants display significantly different opinions from younger participants on the accuracy of the polygraph (Myers et al., 2006). Whatever misperceptions may exist among potential jurors do not appear to be concentrated in any one group of age or education. On the other hand, the insignificant effect of education could prove problematic in correcting any misguided beliefs.

Myers et al. (2006) also researched the respondents' perceived appropriate role for the polygraph in a courtroom. Out of the entire sample, only 4.2% believed that a polygraph could be a singular deciding factor, 62.5% described it as being a useful tool, 25.5% expressed reservations about the practice but conceded a use existed for it, and 7.8% believed it had no usefulness at all (Myers et al., 2006). This is scarcely the overwhelming deference that critics of polygraphs in the courtrooms would accuse jurors of having toward the polygraph. Iacono and Lykken (1997) had earlier surveyed two groups of scientists and experts on the question of polygraphs in the courtroom, and Myers et al. (2006) compared their results to this older survey. In their analysis, they found that the opinions of laypeople and non-professionals differed significantly from the opinions of experts. Myers et al. (2006) found that laypersons had significantly higher opinions of polygraphs (average opinion of 66.7% accuracy) than did experts (average opinion of 44% accuracy). While the difference was significant, it is worth noting that neither rises to the 85% that respondents were informed of before the survey was administered.

In summary, academic and professional opinions of the polygraph are mixed. Despite improvement in examiner training, instrument refinement, and post-test analysis, critics of the polygraph have attacked its potential for error and argued that it still is not accurate or reliable enough to have a place in the justice system (Gallai, 1999). The court history of the polygraph has been dominated by these negative opinions, and it was only recently (i.e., within the last 20 – 25 years) that the instrument found even niche acceptance among the judiciary (*Lee v. Martinez*, 2004). Supporters of the polygraph argue that its potential value outweighs the possibility of error, and that it is no longer

justified to deny the instrument its place due to old case history (Ben-Shakhar et al., 2002). Because of government classification of most polygraph research, there is relatively little information and study on polygraphs in comparison to public opinions on the courts or the justice system in general. This thesis addressed this lack of study and provided a valuable insight into a topic worthy of further research.

Validity of Mock Trial / Jury Research

Actual jury proceedings would naturally be the best way to study juror perceptions and beliefs, but this is not achievable due to the shroud of secrecy that covers most jury proceedings. Juries are prohibited from disclosing or talking about matters that were discussed during deliberation. As such, researchers have used mock trials and juries in their work. Bornstein (1999) explained one significant criticism of this practice, namely the difficulty of creating a realistic simulation for the mock jury to consider. This criticism has come both from outside criminal justice academia as well as from within. Those who have dedicated years of study to the real processes of the justice system are all the more aware of the problems associated with replicating it in a laboratory setting (Bornstein, 1999). A significant quantity of mock trial research has been done with samples composed of undergraduates and college students, which while passable, was not a true representation of the larger public. The format that the research took (e.g., actual staffed trial simulation vs. written summaries) could also impact the validity of conclusions drawn from the research. Ideally, full trial simulations would be the method to employ for researchers, but they are also far more difficult to organize and cost much more in terms of manpower and money. A written trial summary or vignette is much easier to dispense, but will by nature exclude the human element of evidence

presentation, trial discourse such as legal arguments and witness testimony, and jury deliberation; all three topics are important components of real criminal proceedings (Bornstein, 1999).

Bornstein (1999) analyzed 20 years' worth of mock trial / jury research to determine whether simulation quality had improved over time. He found 113 total studies involving mock trial / jury research from 1977 to 1996. Interestingly, Bornstein (1999) found no evidence of quality improvement over time. Of the 113 studies, 65% had undergraduate students play the role of jurors, and 55% used written trial documents. In spite of video equipment becoming more accessible, Bornstein (1999) found no significant increase in the usage of video-based mock trial simulations. If anything, Bornstein (1999) contended that mock trial research was "becoming less realistic over time" (p. 13). In the face of this, he nonetheless argued that disconnect from reality may not matter much. Bornstein (1999) admitted several limitations to mock jury research, chief among them the contemptuous view of some courts if research methods did not closely mirror reality. To counteract this initial skepticism, Bornstein (1999) suggested that researchers seeking to understand jury beliefs first use more academic simulations, using the conclusions to improve and re-adjust the measurements for real-world jury study. Even then, this is no guarantee that the courts would interpret researcher conclusions favorably.

Breau and Brook (2007) also investigated the validity of mock jury research. The newer data set, in comparison to Bornstein (1999), was already an aid to validity, though it was certainly not optimal. To study this topic, Breau and Brook (2007) performed fieldwork involving multiple mock juries. Breau and Brook (2007) were ultimately

attempting to discover the degree to which realism impacted mock juror experience (i.e., if they believed the proceedings to be real, there would be more deliberation) They set up four mock trials at Duke University's law school in 2004, all of which had the premise of a law student having broken the school's honor code (Breau & Brook, 2007).

In the first two cases, the students were instructed and believed that the proceeding was real and the honor code violation was legitimate. In the other two, the students were told that they were participants in a mock jury experiment (Breau & Brook, 2007). To aid in the illusion of realism, in the two 'real' trials, the law student accused of violating the honor code was present in the courtroom as the 'defendant'. In the two overtly mock trials, the individual was not present. The four trials did not run simultaneously, instead running in separate pairs (i.e., both 'real' trials, then both mock trials) at points during the semester. Conclusions from the first 'real' and mock trials were used to improve the quality of the second pair. Before being allowed to deliberate, juries in the two 'real' trials were informed that their decision would be binding and permanent on the 'defendant' student. In contrast, for the overtly mock trials, jurors were instructed that they were participants in a trial simulation measuring the usefulness of the school's honor code and the appropriateness of present sanctions. In all four cases, jurors had reason to take the proceedings with some measure of seriousness, though this effect was certainly more pronounced for the two 'real' trials (Breau & Brook, 2007).

Breau and Brook (2007) saw a demonstrable impact of realism on jury deliberation time and verdict. The two 'real' juries deliberated for 40 and 90 minutes respectively, while both mock juries deliberated for only 30 and 25 minutes respectively. Sanctions were also largely different between 'real' and mock juries. In the first 'real'

case, the jury believed that the proceedings were biased against the ‘defendant’ and refused to impose any sanctions (Breau & Brook, 2007). This stood in sharp contrast to the guilty verdict of the first mock trial, where jurors quickly determined both guilt and appropriate sanctions. For the second ‘real’ case, the deliberations lasted so long that researchers had to cut them off, resulting in a hung jury. In the second mock trial, the jurors once again quickly reached a verdict and devised appropriate sanctions. A post-proceedings questionnaire was administered to all four groups, and the two ‘real’ juries were informed of the nature of the experiment (Breau & Brook, 2007). When asked whether they believed they had put critical, genuine thought into the proceedings, both ‘real’ juries indicated agreement and both mock juries indicated disagreement. In both rounds of trials, the information presented to ‘real’ and mock juries was identical, eliminating procedural difference as a potential source of error. It was transparently and logically obvious that realism had a potent effect on juror decision-making and deliberation (Breau & Brook, 2007). A possible future research idea that Breau and Brook (2007) theorized was that, in response to real crimes, a mock jury be drawn in parallel to the real jury, who would witness the same trial and be instructed identically to the real jury, except for informing the mock jurors of their role as participants in mock trial research. Unfortunately, Breau and Brook (2007) viewed it as very unlikely that many judges would permit their courtrooms to become places for research experiments. In conclusion, the issue of jury secrecy makes real-world study difficult, but it may be worthwhile for researchers to create an illusion of realism to increase the validity of their conclusions, given the impact of realism found by Breau and Brook (2007). However, to create realism, researchers will run into the ethical problems of accusing individuals of

crimes for academic study and deceiving participants into believing they are making rulings that could affect the freedom or rights of another.

In order to avoid these moral issues, this study utilized trial vignettes. A staffed simulation would have been preferable but was deemed unfeasible due to the intrusiveness on selected classes and the monetary cost. For future study, particularly non-student samples, a mock trial would be preferable. It was accepted that the trial vignette lacked the illusion of reality and deliberations. However, layperson opinion of the polygraph is grossly understudied relative to most other topics in the justice system, and as such, any contribution the field is worthwhile. In order to measure these perceptions, a survey was created that measured a wide range of variables all related to the central question: what do jurors believe regarding the polygraph? It is possible that conclusions drawn here could influence policy on polygraph inclusion in some jurisdictions.

CHAPTER THREE: METHODOLOGY

The ultimate goal of this research was to determine the attitudes held by Boise State students on the use of polygraphs in the criminal justice system. Measuring attitudes inevitably ran into concerns about validity and reliability. It was important to ensure that the measures being utilized both cover the topic being surveyed and that the questions were understood in mostly the same way across respondents. Since the aim of this research was to measure perceptions on the use and accuracy of the polygraph, the measurement choice seemed rather simple: a survey. However, in order for this survey (and any surveys, by extension) to have had any degree of validity or reliability, caution had to be taken to ensure that the questions were operationalized properly.

This study was conducted to examine the comparative weight lent to polygraph evidence (relative to DNA, eyewitness accounts, and other forms of criminal evidence) by potential jurors. Methodologically, it was a modified replication of the study by Myers et al. (2006), which analyzed the same broader question. This is an important topic to research, especially given the present state of available research on polygraph perceptions.

For the purposes of this study, a quantitative survey was administered to students of Boise State University, which aimed to discover the opinions and beliefs of students regarding the polygraph, its accuracy, and its usage in the justice system. In forecasting responses, it was imperative to understand that students' attitudes could be shaped by two potential sources of error: tradition and authority (Milgram & Zimbardo, 2013). Those

who have no experience or in-depth knowledge of the legal system could have incorrect or incomplete information or assumptions regarding polygraph accuracy or police usage of the technology. Similarly, if a respondent had taken a polygraph, they would have likely spoken to the examiner, who likely would not have given them a comprehensive summary of present literature on the polygraph due to the impracticality of condensing so much information into a conversation. Human reactions and ‘reading’ physical cues are not a well-known science among the masses, especially when it comes to interpreting responses (O’Toole & Bowman, 2012).

Research Questions and Hypotheses

To reiterate, the central research question to be answered was: what are the general attitudes of potential jurors towards polygraph evidence? There were several hypotheses to be tested in this study:

H₁: Potential jurors will have only moderate faith in the polygraph, ranking it somewhere in the middle when other types of evidence are presented.

H₂: Potential jurors who are in criminal justice courses are likely to view the polygraph more skeptically than those who are in non-criminal justice courses.

H₃: Potential jurors who receive a brief synopsis (less than one page) of the polygraph literature in addition to the vignette / questionnaire will view the polygraph more skeptically than those who do not receive this pre-test information.

In summary, the expected outcomes were very similar to those of Myers et al. (2006), that is, there is no reason to believe that the polygraph ‘usurps the role of the jury’ as some justice professionals fear.

Conceptual Definitions

To answer the research question clearly and test the hypotheses with clarity, accurate and current conceptual definitions were provided to respondents. There is not an official government definition for *polygraph*, but using a reputable dictionary definition seemed an appropriate alternative. While this may seem to some to be unnecessary, it was important that those who participated in this study did not immediately jump to the assumption that the polygraph instrument is a lie detector, as this would have been a threat to validity. The Oxford Dictionary does not explicitly equate ‘polygraph’ to ‘lie detector’. Identifying the importance of both instrument and examiner for survey respondents was important.

Distinguishing types of evidence was also important, since respondents to the survey were asked to rank them in order of strength and credence. It was important to distinguish between direct and circumstantial evidence. The Duhaime Legal Dictionary defined *direct evidence* as follows: “evidence tendered at trial in the form of recounting of personal observations, or a document which directly establishes a fact sought to be proven” (Duhaime, 2017, n.p.). Because polygraphs are never administered in real-time (as the crime is occurring), they always fall under the category of *circumstantial evidence*, which the Duhaime Legal Dictionary defined as “evidence which may allow a trial judge or jury to deduce or logically infer a certain fact from other established facts, which have been proven” (Duhaime, 2017, n.p.).

Myers et al. (2006) used four other types of evidence as comparison points for polygraphs: fingerprint evidence, DNA-based evidence, eyewitness evidence, and handwriting analysis. The definitions for each of these were fairly self-evident, though

this study distinguished them specifically for the benefit of respondents. *Fingerprint evidence*, for the purposes of this research, included any police-gathered and analyzed materials which required a professional fingerprint examiner (Myers et al., 2006). Similarly, *DNA evidence* included all forms of blood work and body fluid analysis. *Eyewitness evidence* involved all forms of individual testimony (either the victim or bystanders, excluding the distinct concept of expert witnesses). *Handwriting analysis* was considered police investigation and subjective analysis of handwriting where applicable (Myers et al., 2006).

Study Design, Sampling Procedure, and Data Collection

The sample in this study consisted of nine total classes; five of which were criminal justice courses. The sampling method used was a stratified cluster sample. For the non-CJ courses, one course was drawn from each of the ‘levels’ in the undergraduate catalog (i.e., 100-level, 200-level, 300-level, 400-level). For the CJ courses, three did not require upper-division standing while the other two did require it. This was similarly intended to ensure that the sample was representative of the wider student body. The non-CJ courses were drawn from the university’s Student Center class search function. Beginning alphabetically, every class subject was numbered sequentially from 1 to 122. A random number generator randomly selected one of these course subjects. After the four course subjects were chosen, the search function was used to identify all available class sections under a particular subject by searching for any available course numbered from 100 to 499. After identifying the total number of available courses, these were numbered sequentially, with the lowest course number being labelled as 1. A random number generator then selected a particular class section. For example, if the random

number generator had produced 1 when selecting subjects, it would have chosen the course subject titled Academic. There were a total of 12 class selections under this prefix with course numbers from 100 to 499. These 12 classes would have been numbered, and the random number generator would have chosen one of them.

Upon selection of a specific course (e.g., MATH, 400-level, non-CJ), the following selections could not be in that same category; that is, another selected class was then prohibited from being 400-level, and if the random number generator happened to select another class in this category, the generator was used again (and as many times as necessary) until it selected a course in line with all the criteria. To select the criminal justice courses, the CJ discipline was chosen in the Student Center and all available courses were numbered accordingly. The random number generator was used to select five classes, two of which were upper division-required, and three were not. Classes were excluded if they were held online or were single-day workshops. The exclusions were due to the impracticality of meeting and speaking to the class in-person, an important component of the methodology. An online course was initially chosen in the sample, but upon discovery that it was held online, another course was selected to take its place. The chosen classes were all face-to-face and lecture-based. Drawing classes from each level helped to ensure that the sample included individuals from multiple levels of undergraduate work.

Once the total sample of courses was drawn, an email was sent to each of the professors teaching the respective class. The non-CJ courses were: PHIL 103 (Introduction to Philosophy: Moral Problems), SOCWRK 201 (Foundations of Social Work), BASQ-STD 380 (Colloquium in Basque Studies), and HEP 440 (Health

Promotion Programming). The CJ courses were: CJ 101 (Introduction to Criminal Justice), CJ 102 (Introduction to Police), CJ 103 (Introduction to Law and Justice), CJ 375 (Criminal Procedure), and CJ 426 (Statistics). This email contained a brief introduction of the researcher and the project, an attachment with the text of the informed consent document and the survey, and concluded with a request to attend a class to explain the project in greater detail and allow for any questions to be addressed in-person. The professors were free to decline, as a few did, and if they did, the sampling framework explained above was used to select a replacement course. After obtaining professor approval, the researcher attended the previously-agreed class time and took three to five minutes explaining the study, its purpose and importance, and answering any questions from students. The link was also provided to the professors at this time, and they were asked to send the link out through Blackboard or via class-wide email. This prevented students from having to manually copy down and rewrite the rather long URL to access the survey. It was emphasized in every class that the survey was strictly voluntary and that the professor's provision of the link was not to be taken as any sort of requirement to participate. To further ensure the survey was taken voluntarily, no extra credit or incentive was introduced in any course. It was certainly true that an incentive could have resulted in higher response rates, but the difficulty of working out an acceptable alternative for each class for non-participants was deemed to outweigh the benefits of incentives.

The Qualtrics survey opened on October 9, 2017 and was available until November 3, 2017. The relatively lengthy amount of time allowed students to take the survey at their convenience and that those who did respond had enough time to give each

question appropriate thought, thus increasing validity. Furthermore, leaving the survey as purely optional also avoids any issue of coercion or socially-influenced responses. The survey took approximately 10 – 15 minutes to complete and was 17 questions in length. First, it included a page explaining informed consent in detail, which provided the respondent with references to counseling services if necessary and contact information for the researcher if they had any further questions before or after completing the survey (Appendix A). The first page also explained precisely that users could exit the survey at any time if it caused them distress, and that should they do this, no penalty would be incurred. Clicking the ‘next’ arrow from this first page was an indication of informed consent and brought respondents to the page of relevant definitions (Appendix B). The trial vignette followed the page of definitions (Appendix C). After proceeding through the vignette, half of all respondents were then provided an additional narrative summary of polygraph literature for their reference to be used in the last question of the survey. Respondents who received the additional information were expressly notified of their selection. For the respondents who were not selected to receive this additional information, the survey clearly stated that they had not been selected by a page which stated, “You were not randomly selected to receive an additional literature summary”. After the two groups were determined, the 16 main survey questions were administered. Question 17 pertained to the additional information, and, for this, the selected respondents answered on a continuum. Question 17 was also present in non-selected group surveys, but respondents here used the answer option that indicated they had not received the information. There was only one version of the survey and both groups of respondents received the same documentation. After completing the survey questions, an

option was provided for respondents to offer their comments or criticism of the project and its implementation. This allows for further improvement of the survey should it be administered again in the future.

This form of questionnaire was chosen over interviews and observations for several reasons. Involving real jurors was not feasible given that all known jurisdictions keep jury deliberations strictly secret, and in certain cases, jurors may be prohibited from discussing the case even after a verdict has been reached. The unpredictable length of jury deliberations would be an issue if real cases were to be used. Inclusion of real, ongoing cases would place the researcher in a dangerous place of potentially serving as an impetus for different behavior if the jurors knew that their behavior and decisions were being used in an academic study. Interviews also posed a risk of allowing the introduction of personal biases and opinions from the interviewer that are not pertinent or helpful in an academic sense. Furthermore, the logistics and resources present for the researcher made conducting interviews or random samples of a larger community impractical.

Dependent Variables

There were multiple dependent variables in this survey. Due to the unavailability of Myers' et al. (2006) methodology, the measurement questions for each were approximated based on their description in the original study. Guilt of the offender was the first dependent variable (see Appendix B, question 1). It was dichotomous, reflective of the range of decisions allowed to U.S. juries. Confidence in the verdict was a second dependent variable and was measured at the ratio level as a percentage (Appendix B, question 2). This allowed respondents to more precisely enumerate the level of

confidence they had in their overall decision. Overall juror confidence in polygraphs was another dependent variable and was measured at the ordinal level by three different questions (Appendix B, questions 3, 7, and 8). Questions 5 and 9 asked respondents to grade their confidence on a five-point scale. The five answers were “not influenced at all”, “slightly influenced”, “moderately influenced”, “greatly influenced”, and “extremely influenced” (Myers et al., 2006, p. 9). Question 9 asked respondents to grade confidence in polygraphs relative to the four other types of comparison evidence listed by Myers et al. (2006). Respondents answered whether polygraph evidence should be considered more or less trustworthy in relation to the other types. All four types of comparison evidence (i.e., fingerprint, DNA, eyewitness, and handwriting) were separate, which allowed respondents to grade polygraphs relative to that individual type of evidence. Support for expert testimony on polygraphs was the penultimate dependent variable and was measured at the nominal level (Appendix B, questions 5 and 6). Question five focused on the use of expert testimony in a positive manner; that is, allowing polygraph experts to testify in court that an individual was being truthful when they denied guilt on the polygraph. Question six, in contrast, analyzed the use of expert testimony in a negative manner; that is, allowing a polygraph expert to testify in court that an accused individual was being deceptive when they denied guilt. To explore H₃, question four asked respondents to grade their own ability to informally detect lies (i.e., through conversation) on a percentage scale from 0 to 100%.

Independent Variables

Due to concerns over anonymity, race was excluded from the list of demographic factors analyzed in the study. The concerns stemmed from the large disparity in minority

enrollment at the university. Per the enrollment data from Fall 2016 (the latest time for which data is available), Whites accounted for almost 75% of all enrolled students. With such a large disparity, it would theoretically be possible to link a particular response to a student, which was not intended. As such, race was excluded from the analysis.

There were two independent variables in this study. Enrollment in a criminal justice class was measured at the nominal level by question 12, which asked respondents to indicate yes or no as to whether the class is listed under the CJ prefix by the Student Center site (Appendix B). Enrollment in a criminal justice course was also selected as an independent variable to gauge whether it led to more critical attitudes towards the polygraph, to test H₂. Receipt of the additional summary and its influence were measured at the ordinal level by question 16, which asked respondents whether they had received the document, and if so, the extent to which it had impacted their decision. For respondents who did not receive the document, an option was provided on this question to indicate that they had not received it

Four control variables were present in the survey: age, sex, class level, and personal / familial involvement in the criminal justice system. Age was measured at the ratio level (Appendix B, question 9). Sex was measured at the nominal level (Appendix B, question 10). Class level was measured at the ordinal level on a four-point scale, with the acceptable answers being 100-level, 200-level, 300-level, and 400-level (Appendix B, question 11). Personal / familial employment in a criminal justice agency was measured at the nominal level, and respondents answered yes or no (Appendix B, question 13). Personal or familial involvement in polygraph or criminal justice work was also selected as a possible control for more supportive attitudes toward the polygraph. Age and sex are

frequent control variables in research, though Myers' et al (2006) analysis did not include them in their discussion. Class level was chosen as a control variable, and was measured at the ordinal level by question 12 (Appendix B).

Validity and Reliability

When dealing with perceptions and opinions, there is always going to be a degree of difficulty in attaining valid, reliable results. This was one reason that this study was partly replicative in nature, and even the baseline study, by Myers et al. (2006), used measures that had been used before their research was conducted. They indicated that similar questions had been used by larger agencies, such as Gallup, in a much earlier (1984) poll to measure attitudes. Each question on the instrument had a reasonable level of face validity. The survey asked individuals to descriptively rate where polygraph evidence stands, which made logical sense in that it would lead to answers which are tied to that question. Furthermore, the provision of relevant definitions and allotment of time before test administration to inform all respondents helped to ensure that they had a solid understanding of what was expected and what their answers signified. The likelihood of poorly-written or unclear questions would have been higher if this study had used newly-invented measures. Myers et al. (2006) indicated that their questions had been drawn from previous opinion surveys administered to experts by Amato (1993), Gallup Organization (1984), and Iacono and Lykken (1997). In that particular study, Iacono and Lykken (1997) did not report any substantial criticism of the validity and reliability of their questions, though it was conceded that their questions were administered to experts rather than students as was the case in this study.

In addition, while no questions were direct repeats of one another, several were directly related: such as the questions which asked how much the polygraph evidence weighed on the decision inside the sentence decision and then outside of a criminal proceeding. Asking these triangulation questions gained a depth of information and helped ensure that respondents' understanding of the questions was correct (Tokunaga, 2016).

The choice of questionnaires / surveys was deliberate, as they are the most common form of formal public opinion research, though they are being challenged by social media data analysis (Assessing New Ways, 2016). Businesses, such as Gallup, are very well-known for conducting public opinion surveys, and they have been referenced in academic publications such as Myers et al. (2006). Not only would interviews have required additional resources that were not presently available, aggregating hundreds of transcribed interviews into legible, interpretable data would have been significantly more time-consuming. Numerical scales also decrease the likelihood of personal bias affecting validity than would researcher interpretation of interview code words (Tokunaga, 2016). In addition to this, the choice of anonymously-administered online surveys was advantageous given that some opinions on polygraph may be considered controversial, and individuals would likely hesitate to express these opinions even in the presence of an academic interview. An anonymous survey allowed these potentially-contentious opinions to surface in comparative safety, avoiding the problem of social desirability bias (Tokunaga, 2016). In line with the standards suggested by the American Association for Public Opinion Research (AAPOR), to increase validity, the questions were replicated from previous research as much as possible. The provision of definitions and clear

delineation of acceptable answers increased internal validity in the survey as well (AAPOR, 2015).

Limitations

As with any research, there were limitations that have to be conceded. The first issue with a project such as this was the difficulty / propriety of condensing the depth of information that emerges in a criminal trial down to a readable, quickly-understandable vignette. It was a valid criticism to argue that it was impossible to condense so much information without losing valuable depth and context. A particular problem was avoiding 'prejudicing' respondents towards positive or negative responses based on the vignette and especially those who received the review of polygraph literature. Precautions were taken to ensure that no bias could be observed in the summary of literature. Four positive court cases / articles in favor of the polygraph were included, as well as four court cases / articles that were critical of the polygraph. Equal space was allotted to both categories.

The research design itself also imposed its own limitations. Because the survey was a single case study of students, it was impossible to say with certainty that the patterns observed would remain true from year to year. Furthermore, because the test was administered only once in the fall semester, it was impossible to account for the potential change in opinion that could occur as classes proceeded and professors engaged in discussion with their classes. Because of this, the research cannot account for the impact of ongoing education on opinions about the polygraph. The results of this survey could only be considered a 'snapshot' of the opinion of the sample at the time (Tokunaga,

2016). Because there were no pre- or post-tests, this precluded any examination of causal relationship between suspected factors (Anderson, 2017).

Second, while questionnaires had a rapid turnaround time when compared to interviews or direct observation, it was possible that even with the goal of keeping documentation brief that survey fatigue occurred, especially so for those who received the additional summary (Tokunaga, 2016). Questionnaires of this sort also did not allow the researcher to clarify misunderstandings or directions after the test had begun, and with the amount of information that was presented in a condensed vignette and optional summary, there was also undeniably a risk of some respondents being unclear or unsure of their answers (Tokunaga, 2016). This made it all the more important to allow sufficient time beforehand to explain in full detail the study's purpose, scope, and future usage of its data. However, with the degree of preparation and rigor that was applied to this study, the risk of confusion seems manageable.

As with any student-based sample, conclusions drawn from this data cannot generalize perfectly towards the larger public, since there were demographic differences between the two. College students, on the whole, are younger and tend to be at least slightly more liberal than does the general public (Farnworth, Longmire, & West, 1998). Druckman and Kam (2009) addressed the concerns that other researchers may have with student samples. External validity is compromised in a student sample if the research is attempting to identify causal relationship between variables, which this study was not intending to do. In a larger, more comprehensive study, it would be useful to have a separate sample drawn from the public to compare to the findings from the student sample.

It was also possible that the classes selected by the random number generator had widely different numbers of students enrolled, which could have greatly skewed the statistical analyses of a particular category. This was the case with the 100-level courses, as they had many more students than the 400-level classes. The possibility of this occurring was also a threat to validity, as it would have been statistically invalid to attempt to draw conclusions if the sample size for one category of class was only 30 compared to another category's 200. Any discrepancies by sex or age, compared to the proportions of Boise State's student population as a whole (per the Fall 2016 Census Enrollment) were noted in the data analysis section. The choice to take a stratified cluster sample of each class level was intended as an aid to generalizability. If a stratified cluster sample had not been taken, it was possible that the sample could have been disproportionately comprised of members of one class level, which would have diminished its generalizability to the larger student sample.

To determine whether the hypotheses were supported, chi-square analyses were conducted. Because most of the variables were nominal or ordinal, this prevented stronger statistical analyses from being performed. Even descriptive statistics were limited by ordinal data, as the mean, standard deviation, and coefficient of variation do not produce meaningful results with ordinal variables. As such, chi-squares were the best available measure to determine whether relationships existed between the dependent and independent variables.

CHAPTER FOUR: RESULTS

Sample Descriptives

Despite attempting to obtain a large, broadly representative sample with the methods chosen, it was evident that this was unfulfilled. The only control variable in agreement with previous enrollment data was that of age, with a median value of 21 years. Aside from this, choosing multiple levels of courses clearly did not serve its intended purpose. With over 70% of all responses coming from 100-level students, it was difficult to suggest with much confidence that this was representative of the attitudes of any other class-level (Table 1.0). However, with age also being similarly concentrated at the lower end, this was less surprising of a result. The mean was skewed upwards due to the presence of several older respondents (above 60 years old). Also somewhat expectedly, more responses were from CJ students than from non-CJ students. This made some sense, as students in CJ courses were more likely to have their professors encourage them to participate or to feel it was worthwhile to answer such surveys, since it was related to what they are studying. However, because the groups were not equal, conclusions were less generalizable outside the criminal justice discipline.

Because the course was forwarded through class emails or posted on BlackBoard, there was no definitive value for the number of students that could have received the link (i.e., not all students will check emails or Blackboard). However, based on total enrollment data from each class, there were roughly 346 possible recipients of the email. Only 113 opened the survey link at all. Out of the 113 opened survey links, there were

only 56 completed surveys. In terms of response rate, using the total number of opened survey links (113), the response rate was approximately 33%. Since 56 of these 113 total links were actually completed, the completion rate was roughly 50%. The combination of young age and low class-level made the very small sample of upper-division required students very much expected. Out of all respondents, only four were students in an upper division-required CJ course.

Table 1.0 Sample descriptives

<i>Variable</i>	<i>n</i>	<i>Percentage</i>	<i>Mean</i>	<i>Median</i>	<i>SD</i>
Sex					
Male	16	28.6%			
Female	39	69.6%			
Missing	1	1.8%			
Course Level					
100 Level	40	70.2%			
200 Level	2	3.5%			
300 Level	9	15.8%			
400 Level	5	8.8%			
In a CJ Course					
Yes	39	68.4%			
No	15	26.3%			
Missing	3	5.3%			
Upper Division CJ Course					
Yes	5	8.8%			
No	29	50.9%			
Not Applicable	23	40.4%			
Employed in CJS					
Yes	4	7.0%			
No	50	87.7%			
Unsure	1	1.8%			
Age	56		25.89	21	12.68
Verdict					
Guilty	20	35.1%			
Not Guilty	37	64.9%			
% confidence in verdict	57		57.92%	51%	21.76
Believed % accuracy of polygraph	57		54.80%	51%	20.84
Personal detection accuracy %	56		50.1%	50%	20.36

<i>Influence of vignette</i>		
Not influenced	8	14.3%
Slightly	26	46.4%
Moderately	16	28.6%
Greatly	6	10.7%
Support examiner testimony of non-deceptive (+) exam		
Yes	31	55.4%
No	25	44.6%
Support examiner testimony of deceptive (-) exam		
Yes	33	58.9%
No	23	41.1%
General opinion of polygraph		
Only evidence needed	1	1.8%
Useful when corroborated	24	42.9%
Questionable but can be used	25	44.6%
Not useful	6	10.7%
Polygraph v. fingerprints – trustworthiness		
Polygraph more	8	14.3%
Polygraph less	48	85.7%
Polygraph v. DNA – trustworthiness		
Polygraph more	7	12.5%
Polygraph less	49	87.5%
Polygraph v. handwriting – trustworthiness		
Polygraph more	21	37.5%
Polygraph less	35	62.5%
Polygraph v. eyewitness – trustworthiness		
Polygraph more	15	26.8%
Polygraph less	41	73.2%
Influence of extra literature summary		
No influence	3	5.8%
Slightly	7	13.5%

Moderately	9	17.3%
Greatly	2	3.8%
Did not receive	31	59.6%

Bivariate Analyses

Due to the ordinal level of measurement for most dependent variables, chi-square analyses were the best-suited statistical measure to identify relationships between variables. Results of the verdict were divided, as only 35.1% found the accused guilty. Percentage confidence in the verdict itself was relatively moderate, with a mean of 57.9% (Table 1.0). H₁ theorized that respondents would have only moderate faith in the polygraph, rather than the unquestioning acceptance that previous authors have written. Based on the data in Tables 1.0 and 1.3, it appeared safe to say that H₁ was supported. When compared to the other four types of evidence (fingerprint, DNA, handwriting, eyewitness) used in the Myers et al. (2006) study, polygraphs were seen as less trustworthy than all of them. Average confidence in the polygraph was also relatively low at 54.8%. This result was indicative that the instrument did not ‘invade the province of the jury’ (Pettit, 2007). Respondents also clearly showed that the provision of polygraph evidence within the trial vignette did not have a very strong impact on their overall verdict. The median value for influence of the vignette evidence was 2, which corresponded to ‘slightly influenced’.

H₁ was further strengthened by the surprisingly skeptical attitudes towards examiner clarification of results. As seen in Table 1.1 below, only 55.4% of respondents believed that examiners should be permitted to explain a non-deceptive result. The reasoning for such skeptical attitudes toward even clarification would be an interesting continuation of this research. Distrust in the polygraph also extended to examiner

testimony regarding a deceptive (i.e., a failed test) result, as seen by Table 1.2.

Respondents did not believe that examiners should be allowed in court even to clarify that a suspect had failed a polygraph exam. This suggested that respondents had much more nuanced views of the polygraph than previous studies have shown.

Table 1.1 Support for examiner explanation of non-deceptive results

<i>Response</i>	<i>n</i>	<i>Response Percentage</i>
Yes	31	55.4%
No	25	44.6%

Table 1.2 Support for examiner explanation of deceptive results

<i>Response</i>	<i>n</i>	<i>Response Percentage</i>
Yes	33	58.9%
No	23	41.1%

H₁ was also supported by respondents' opinions on the polygraph's role and function in a courtroom. As Table 1.3 shows, only one individual believed the polygraph could entirely decide a case. This is the sort of opinion that, if widespread, would lend credence to the 'invading the province' thought. It was very clear from the numbers in this sample that such thinking was rare. With a modal value that corresponded to "questionable use, entitled to little support against other evidence", support for the polygraph in this sample was marginal.

Table 1.3 General opinion of the polygraph

<i>Opinion</i>	<i>n</i>	<i>Percentage</i>
Only evidence needed	1	1.8%
Useful when corroborated	24	42.9%
Questionable but can be useful	25	44.6%
Not useful	6	10.7%
Total	56	100%

Interestingly, respondents within this sample also appeared to have a far more realistic idea of their personal accuracy at detecting lies in conversation. The self-

reported accuracy mean was lower than that of the polygraph at 50.1% (Table 1.4). This was in line with research on interpersonal lie detection, which suggested that even trained and experienced individuals still did not have lie detection abilities significantly above that of a layperson, who themselves were not very skilled at detecting lies in conversation. It was intriguing to note that the averages for polygraphs and persons were not statistically significantly different.

Table 1.4 Perceived accuracy of the polygraph

<i>Respondent Perception</i>	<i>n</i>	<i>Mean</i>	<i>Median</i>	<i>Mode</i>	<i>SD</i>
Polygraph accuracy	57	54.8%	51%	50%	20.84
Personal detection accuracy	56	50.1%	50%	50%	20.36

Sample respondents did not believe they were particularly adept at detecting lies, but they also did not see much promise in instrumental detection either. However, with both personal accuracy and instrument accuracy, the standard deviation was particularly large compared to the mean itself, with an SD of 20.84 for the instrument and 20.36 for personal lie detection (Table 1.4). The large standard deviation in this instance weakened the ability to make definitive conclusions (i.e., it was not an argumentative or research-interesting point to state that laypersons had anywhere between 30 to 70% confidence in the polygraph to detect lies). The 30 to 70% range was obtained by taking the mean and shifting one standard deviations' worth of distance in both a positive and negative direction.

The second hypothesis argued that criminal justice students would be more apt to view the polygraph in a critical light compared to those not enrolled in such courses. To analyze this, chi-square analyses were conducted to determine if there was a statistically significant relationship between the aforementioned attribute and their attitudes toward the polygraph. An alpha of .05 was chosen as the threshold for significance in all chi-

square tests. None of the chi-square tests resulted in significance for respondents enrolled in criminal justice courses.

Table 1.5 Chi-square analyses CJ course with dependent variables

<i>CJ Course Comparisons</i>	<i>n</i>	<i>X²</i>	<i>Degrees of freedom</i>	<i>Significance (2-sided)</i>
Verdict	57	.036	2	.982
Graded verdict	57	48.387	2	.617
Confidence in polygraph	57	50.301	56	.689
Influence of evidence	56	1.646	6	.949
Clarifying non-deceptive	56	.066	2	.968
Clarifying deceptive	56	.539	2	.764
General opinion of polygraph	56	3.671	6	.721
Polygraph vs. fingerprint	56	5.346	2	.069
Polygraph vs. DNA	56	4.167	2	.124
Polygraph vs. handwriting	56	1.255	2	.534
Polygraph vs. eyewitness	56	.576	2	.750

While the small sample size weakened the predictive ability of bivariate correlations, they were also conducted to explore any further relationships between CJ course students and their attitudes toward the polygraph. Additionally, low cell counts were an issue for every chi-square analysis between CJ course students and their answers. The same alpha of 0.05 was used as the standard for significance. Ultimately, the chi-square results also were indicative of a lack of relationship between variables. The only variables to have any significant correlation at the .05 level with CJ course students were other control variables: namely, whether the course required upper division standing and whether the individual had any personal or familial ties to polygraph work or criminal justice agencies. The conclusions from these bivariate correlations also did not suggest any unusual relationship: it was not an argumentative point to suggest that respondents in CJ courses were more likely to be members of upper-division required classes (Table 1.6). Due to the complete lack of statistically significant results H_2 was not supported.

Table 1.6 Bivariate correlations

	<i>Respondent in a CJ course</i>
Upper division requirement in CJ course	.300 (.023) n=57
Personal or familial involvement in CJ	.357 (.007) n=56
Polygraph vs. DNA evidence	-.222 (.100) n=56

The final hypothesis argued that those randomly selected to receive the one-page literature summary would have significantly harsher views on the polygraph than those who did not. Identically to H₂, chi-square analyses were conducted to observe any significant effects that the extra information had. Unfortunately, low cell counts were once again present for all chi-square analyses in terms of the extra literature summary. One relationship was statistically significant. For those who received the extra literature summary, there was a significantly different influence of the polygraph evidence in the trial vignette, with an alpha of .002. Respondents who received the extra literature summary had more critical views of the polygraph than those who did not receive it. This suggested that the additional information was actually read by the respondents. It was notable that the influence of this additional information was negative, even though equal space and sources were provided for both supportive and critical literature. Problematically, low cell count was an extreme issue here, as 90% of cells for this question had an expected count less than five. Conceptually, this made some sense, as it was likely that having new information (both positive and negative) regarding the polygraph could have changed its swaying power compared to those who did not have the information. Bivariate correlation using Pearson's R was also conducted, though it was

further weakened by the extremely small sample size for this question: only 21 out of the total sample of 57 received the extra document. For the bivariate analysis, an alpha of .05 was used. In this analysis, none of the bivariate comparisons reached significance.

Because only one chi-square relationship was found to be significant between receipt of the literature summary and attitudes towards the polygraph, it appeared that H₃ was only partially supported. The literature summary did not significantly impact any other relevant responses.

Table 1.7 Chi-square analyses literature summary with dependent variables

<i>Literature Summary Comparisons</i>	<i>n</i>	<i>X²</i>	<i>Degrees of freedom</i>	<i>Significance (2-sided)</i>
Verdict	52	3.252	4	.517
Graded verdict	52	80.768	100	.921
Confidence in polygraph	52	120.810	104	.124
Influence of evidence	52	31.332	12	.002
Clarifying non-deceptive	52	4.797	4	.309
Clarifying deceptive	52	1.527	4	.822
General opinion of polygraph	52	10.526	12	.570
Polygraph vs. fingerprint	52	6.033	4	.197
Polygraph vs. DNA	52	4.449	4	.349
Polygraph vs. handwriting	52	3.931	4	.415
Polygraph vs. eyewitness	52	1.906	4	.753

Table 1.8 Bivariate correlations for literature summary

	<i>Literature Summary</i>
Age	-.213 (.130) n=52
Sex	.113 (.426) n=52
Course Level	-.056 (.694) n=52
Upper Division CJ course	.099 (.486) n=52
Personal / familial involvement in CJ	.113 .424 n=52

Discussion

In terms of policy implications, this study's impact was limited, as the sample was composed of college students and therefore was not guaranteed to have been representative of actual public opinion. However, there were several conclusions that could be drawn despite these issues. It was evident that there was little reason for the courts to believe that the polygraph 'invades the province of the jury' (Pettit, 2007, p. 10). With less than half of the total sample having chosen to convict the offender and average confidence in the polygraph at approximately 54%, it was clear that sample respondents were not blankly accepting polygraph results as proof of guilt. Respondents were also divided on whether polygraph examiners should be permitted in court to explain the meaning of certain results, which was unexpected. Potential jurors within this sample absolutely did not display deference or confidence levels that would have suggested they allow the polygraph to 'invade the province', but neither did they completely reject it (Pettit, 2007, p. 10). Similar to the conclusions from Myers et al.

(2006), average opinion of the utility of the instrument was moderately supportive. Potential jurors saw some use for the polygraph, but most indicated that results needed to be corroborated by other evidence. This suggested that the jury was still fulfilling its role as ‘trier of fact’ in the justice system when it came to polygraph evidence. Courts which exclude the instrument out of concern for jury integrity should reconsider whether such beliefs are substantiated by current data, as this study and Myers et al. (2006) indicated that there was very little reason for such consternation. However, it was conceded that court exclusion on the basis of a *Daubert* failure could not be challenged by the results of this study, since the study focused on potential juror perceptions, as opposed to polygraph science itself.

H₂ inquired whether respondents in criminal justice courses viewed the polygraph more skeptically than those not in such courses. Chi-square analyses of questions 13 (whether the respondent took the survey in a CJ course), 1 (verdict), 6 and 7 (support for examiner clarification of results) consistently demonstrated statistical non-significance. Unfortunately, due to the extremely small sample size, it was difficult to say that these results were very representative of any wider population, even just the student body. Nonetheless, H₂ must be rejected, as none of the available statistical measures supported it. It is possible that, in a larger sample or one that includes criminal justice professionals as well as non-CJ workers, a similar hypothesis could be supported. Judges and lawyers who know the precedent and history of the polygraph would be unlikely to support it, where those with a less informed view might be more prone to believe in the polygraph. Enrollment and participation in a CJ course did not have a statistically significant impact on whether attitudes regarding whether or not examiners should be allowed to explain

what the results truly mean. Given that the majority of the sample viewed this practice negatively, it was not surprising that this is the case.

H₃ supposed that individuals who were randomly selected to receive the extra vignette of polygraph literature (with equal inclusion of supportive and critical information) would view the polygraph more skeptically than those who did not receive it. Chi-square analysis were performed with this question (question 16) and the questions from H₂ (verdict and support for examiner clarification), and once again, none of the results were statistically significant. This result was particularly limited by an even smaller sample size, as only 21 respondents received and responded to this final question. With such a miniscule sample, the author was precluded from any larger generalization outside the sample. None of the demographic control variables were found to be statistically significant on the impact of this vignette, which was expected. It was probable that respondents to this survey already had a fairly balanced view of polygraphs, regardless of the information provided. This was a positive conclusion for the sample, but was not supportive of the research questions.

It was possible that the CSI effect was occurring in this study. Because polygraphs are not 'traditional' forensic evidence (e.g., blood spatter, gunshot residue, DNA analysis), jurors may have been less likely to consider its importance in a case when presented with multiple types of evidence. The lack of a federal rule for polygraphs and the resulting different standards among the states also meant that results and policy implications from this study could not be generalized outside of Idaho. Expanded replication (i.e., taking a truly public sample) in other jurisdictions would rectify this lack of information. Further scientific analysis of polygraph theory and technique is also

needed in order for expanded use, as it would have a strong impact on the instrument's ability to pass *Daubert*, which determines whether it is allowed in the courtroom at all. The shaky legal precedent for polygraphs is set in stone, but it would also be important for courts and judges challenge case history if later studies shine a more supportive light on the instrument.

It will be similarly vital for police departments and other state agencies to understand that even course instruction on informal / conversational lie detection does not grant them any sort of sixth sense or make them significantly better than the larger public. This could influence the hiring and teaching practices of these same agencies when it comes to lie detection. Decisions made here could impact the justice system in other areas as well. For example, changes in investigative practices or interrogation methods could easily influence a case as it proceeds through the justice system. It will be imperative for organizations to make decisions that are supported by up-to-date research and appropriate methods. Educating the public on the bleak reality of conversational lie detection would be valuable as well, as it is possible (though not guaranteed) that jurors may engage in their own form of lie detection in a trial proceeding, such as watching a defendant during their testimony. However, it must be conceded that whether jurors actually engage in this behavior was outside the scope of this study.

Different methods of lie detection exist, and it is important for them to be studied and related to the polygraph. Converus is one of these companies, and they have conducted lie detection using instrumentation that tracks eye movements. Proponents of Converus have indicated that it can have a very high (approximately 86%) accuracy rate, and that proper techniques of the instrument could be easily taught to new users (Handler,

2016). The one major drawback at this time would be the rather steep cost of each exam; Converus' EyeDetect machines cost approximately \$4,000 each. Because the criminal process already is rather time-consuming and expensive, it is unlikely that Converus will see widespread implementation any time soon. Regardless, the potential corroborative value of alternative methods of lie detection is very high and will remain a worthy research topic as it continues to be explored (Handler, 2016).

Suggestions for Future Research

For future research, validity could be improved by making the trial and jury proceedings into a full simulation, rather than distilling the facts into a short vignette. It is true that such methodological changes are more personnel-intensive and expensive, but they provide a level of similarity to real experience that is much closer than written summaries and surveys can achieve. Courts would be more likely to accept findings from studies in which respondents' experiences were close to reality. These validity gains should be weighed against the potential for ethical problems in approximating real trials. It is likely that future research conducted in this manner would have to avoid simulating the most serious of crimes, such as murder and rape, due to the obvious ethical issues of accusing an individual of egregious crimes and deliberately leading participants to think that they are ruling on real crimes all for the sake of research. If a vignette-based study were to be used, it would be imperative that the language and reading level be tailored appropriately. Surveys that included with graduate students could use more technical terms than could be used in a study targeting freshmen.

Improvements could also be made by obtaining and referring to the most current policies on polygraph administration and evidence when designing a future study. It

would be of little worth for participants to work with obsolete or unused practices. As such, relevant-irrelevant tests should not be utilized in any future studies that analyze perceptions or opinions, since the National Research Council (2003) indicated that relevant-irrelevant polygraphs were viewed skeptically even among the polygraph community itself. While it would not likely be feasible to administer a mock test to educate all potential jurors on techniques and theory, it would be important to ensure that studies utilize methods that are permitted in judicial settings. Similarly, definitions must be updated and changed when necessary to ensure that they are as close to reality and practice as possible. For studies that include college students as part of the sample, it would be advantageous to look at an additional sample drawn from the public. This would allow for comparison and highlight any differences between the two groups beyond what is already known (i.e., that college students are younger and are more liberal). It would be imperative that student samples were drawn from multiple departments and disciplines within a university, as this would aid in generalizability towards the entire student body.

To expand the depth of information in a future study, qualitative interviews with respondents and polygraph professionals would be a strong suggestion. The addition of a post-closure class visit to the methodology of this study would be one way of increasing available information. Researchers could go to the same classes as before and ask the respondents if they had any further questions pertaining to the survey. This does not guarantee that respondents would provide any further information, but the opportunity to do so would be present, as well as allowing any misunderstandings or vagueness to be clarified immediately. In the current study, no respondent used the provided email to

contact the researcher, so it was possible that the survey did not have any such issues. Again, this would likely increase the cost of the study, but as Breau and Brook (2007) and Bornstein (1999) both concluded, there was a validity increase when depth of information was improved, for participants and researchers alike. In studies that analyze the theory and scientific accuracy of the polygraph, comparison groups from alternative methods (such as Converus) would be greatly useful.

Conclusion

As a modified replication of Myers et al. (2006), their definitions and terms were used to create the framework of this study. H₁ was found to be statistically supported; jurors displayed a moderate level of belief in polygraph accuracy, not unconditional acceptance or anything that suggested they allowed the instrument to fulfill their Supreme Court-defined role as trier of fact. H₂ was not statistically supported. Survey participants in criminal justice courses displayed no significant differences in opinion from respondents outside such courses. H₃ was also not supported by statistical analysis. Receiving the extra information on polygraphs did not result in significantly different opinions on the polygraph's use or accuracy. An overarching theme of the analysis was the problem of small sample size. Originally, logistic regression analysis was planned, but the issue of sample size precluded this from being used. The issue of sample size also increased the likelihood of type II error. Because the sample was very small, it is difficult to conclude that there was even enough data to make a valid decision on the research hypotheses. However, the understudied nature of the topic (polygraph perceptions among laypersons) made this study a valuable contribution to this body of research.

REFERENCES

- AAPOR (2015). *Code of Professional Ethics and Practices*.
- Amato, S. L. (1993). *A survey of the Society for Psychophysiological Research regarding the polygraph: Opinions and interpretations*. Unpublished master's thesis, University of North Dakota.
- Amato-Henderson, S., Honts, C. R., & Plaud, J. J. (1996). *Effects of misinformation on the concealed knowledge test*. PsycEXTRA Dataset. doi:10.1037/e526132012-065.
- American Psychological Association (2004). *The truth about lie detectors (aka polygraph tests)*. (n.d.). Retrieved February 24, 2018, from <http://www.apa.org/research/action/polygraph.aspx>.
- Anderson, B. (2017). *Establishing causality* [Web log post]. Retrieved from <http://info.umkc.edu/drbanderson/establishing-causality/>.
- Assessing new ways to measure public opinion*. (2016, March 1). Retrieved May 04, 2017, from <http://home.isr.umich.edu/releases/assessing-new-ways-measure-public-opinion/>.
- Ben-Shakhar, G., Bar-Hillel, M., & Kremnitzer, M. (2002). Trial by polygraph: Reconsidering the use of the guilty knowledge technique in court. *Law and Human Behavior*, 26(5), 527-541. doi:10.1023/a:1020204005730.
- Blackstone, W. (1778). *Commentaries on the laws of England*. Oxford: Clarendon Press.
- Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior*, 23(1), 75-91. doi:10.1023/a:1022326807441.
- Breau, D., & Brook, B. (2007). 'Mock' mock juries: A field experiment on the ecological validity of jury simulations. *Law & Psychology Review*, 31, 77-92.

- City of Wahpeton v. Skoog*, 295 N.W.2d 313 (N.D. 1980).
- Claar v. Burlington Northern Railroad Company*. 29 F.3d 499 (1994).
- Costelloe, M., Chiricos, T., & Gertz, M. (2007). Punitive attitudes toward criminals. *Punishment and Society*, 11(1), 25-49. doi:10.1177/1462474508098131
- Daubert v. Merrell Dow Pharmaceuticals, Inc.* 509 US 579 (1993).
- Druckman, J. N., & Kam, C. D. (2011). Students as experimental participants: A defense of the 'narrow data base. In J. N. Druckman, D. P. Green, J. H. Kuklinski, & A. Lupia (Eds.), *Cambridge handbook of experimental political science* New York: Cambridge University Press.
- Faigman, D., Fienberg, S., & Stern, P. (2003). The limits of the polygraph. *Issues in Science and Technology*, 20(1), 40.
- The Duhaime legal dictionary*, 2017. Retrieved April 14, 2017, from <http://www.duhaime.org/LegalDictionary.aspx>.
- Farnworth, M., Longmire, D. R., & West, V. M. (1998). College students' views on criminal justice. *Journal of Criminal Justice Education*, 9(1), 39-57,178-180.
- Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923).
- Gallai, D. (1999). Polygraph evidence in federal courts: Should it be admissible? *American Criminal Law Review*, 36(1), 87.
- Gallup Organization. (1984). Survey of members of the American Society for Psychophysiological Research concerning their opinion of polygraph test interpretation. *Polygraph*, 13, 153–165.
- General Electric Company v. Joiner*. 522 US 136 (1997).
- Goodfellow, M., & Kilgore, C. (2013). DUI offenders' beliefs about DUI statutes and DUI law enforcement. *Journal of Drug Issues*, 44(3), 269-280. doi:10.1177/0022042613500052

- Greely, H. T., & Illes, J. (2007). Neuroscience-based lie detection: The urgent need for regulation. *American Journal of Law & Medicine*, 33(2-3), 377-431.
doi:10.1177/009885880703300211
- Han, Y. (2016). Deception detection techniques using polygraph in trials: Current status and social scientific evidence. *Contemporary Readings in Law and Social Justice*, 8(2), 115-147.
- Handler, M. (March 2016). Low base rate screening survival analysis & successive hurdles. *Journal of the American Association of Police Polygraphists*, (1), 31-38.
- Handler, M., Honts, C. R., Krapohl, D. J., Nelson, R., & Griffin, S. (2009). Integration of pre-employment polygraph screening into the police selection process. *Journal of Police and Criminal Psychology*, 24(2), 69-86. doi:10.1007/s11896-009-9050-2.
- Honts, C. R. (2014). Countermeasures and credibility assessment. *Credibility Assessment*, 131-158. doi:10.1016/b978-0-12-394433-7.00004-x.
- Honts, C. R. (2013). *Current FBI polygraph / interrogation practices puts the actually innocent at risk of false confession*. PsycEXTRA Dataset.
doi:10.1037/e571212013-396.
- Honts, C. R. (1997, June 1). *Is it time to reject the friendly polygraph examiner hypothesis (FPEH)?* Retrieved from
<http://truth.charleshontsphd.com/Polygraph/fpeh.html>.
- Honts, C., Amato, S., & Gordon, A. (2001). Effects of spontaneous countermeasures used against the comparison question test. *Polygraph*, 30(1), 1-9.
- Honts, C. R., & Perry, M. V. (1992). Polygraph admissibility: Changes and challenges. *Law and Human Behavior*, 16(3), 357-379. doi:10.1007/bf01044774.
- Honts, C. R., & Quick, B. D. (1995). The polygraph in 1995: Progress in science and the law. *North Dakota Law Review*, 71(4), 987-1019.
- Honts, C. R., & Raskin, D. C. (1988). A field study of the validity of the directed lie control question. *Journal of Police Science and Administration*, 16, 56-61.

- Honts, C., Raskin, D., & Kircher, J. (1987). Effects of physical countermeasures and their electromyographic detection during polygraph tests for deception. *Journal of Psychophysiology, 1*, 241-247.
- Honts, C., Raskin, D., Kircher, J., & Hodes, R. L. (1988). Effects of spontaneous countermeasures on the physiological detection of deception. *Journal of Police Science and Administration, 16*, 91-94.
- Hurwitz, J., & Peffley, M. (2005). Explaining the great racial divide: Perceptions of fairness in the U.S. criminal justice system. *The Journal of Politics, 67*(3), 762-783. doi:10.1111/j.1468-2508.2005.00338.x
- Iacono, W. (2008). Effective policing: Understanding how polygraph tests work and are used. *Criminal Justice and Behavior, 35*(10), 1295-1308.
- Iacono, W. G., & Lykken, D. T. (1997). The validity of the lie detector: Two surveys of scientific opinion. *Journal of Applied Psychology, 82*(3), 426-433.
- Ibek, A. (2012). Methodological problems of polygraph examination. *Internal Security, 4*(2), 175-191.
- Institute for Credentialing Excellence. (n.d.). Retrieved April 12, 2017, from <http://www.credentialingexcellence.org/>.
- Katz, E. (1997). American federalism, past, present and future. *Issues of Democracy, 2*(2).
- Kaukinen, C., & Colavecchia, S. (1999). Public perceptions of the courts: An examination of attitudes toward the treatment of victims and accused. *Canadian Journal of Criminology, 41*(3), 365-384.
- Kozinski, A. (2015). Criminal Law 2.0. *Georgetown Law Journal: Annual Review of Criminal Procedure, 44*(1), 3-44.
- Lee v. Martinez*. 136 N.M. 166, 96 P.3d 291, 2004.
- Marcus, P. (2007). Capital punishment in the United States, and beyond. *Faculty Publications: William and Mary Law School, 61*, 837-872.

- Maschke, G., & Scalabrini, G. (2005). *The lie behind the lie detector* (4th ed.). Retrieved from <https://antipolygraph.org/lie-behind-the-lie-detector.pdf>.
- McCorkle, R. (1993). Research note: Punish and rehabilitate? Public attitudes toward six common crimes. *Crime and Delinquency*, *39*(2), 240-252.
- Milgram, S., & Zimbardo, P. G. (2013). *Obedience to authority: An experimental view*. London: Printer & Martin.
- Military Rule 707. 64 Stat. 109, 10 U.S.C. §§ 801–946
- Myers, B., Latter, R., & Abdollahi-Arena, K. (2006). The court of public opinion: Lay perceptions of polygraph testing. *Law and Human Behavior*, *30*(3), 509-523. doi:10.1007/s10979-006-9041-0.
- National Research Council (2003). *The polygraph and lie detection. Committee to review the scientific evidence on the polygraph. Division of Behavioral and Social Sciences and Education*. Washington, DC: The National Academies Press.
- Offe, H., & Offe, S. (2007). The comparison question test: Does it work and if so how? *Law and Human Behavior*, *31*(3), 291-303. doi:<http://dx.doi.org.libproxy.boisestate.edu/10.1007/s10979-006-9059-3>.
- O'Toole, M. E., & Bowman, A. (2012). *Dangerous instincts: How gut feelings betray us*. New York: Plume.
- Patrick, C. J., & Iacono, W. G. (1991). Validity of the control question polygraph test: The problem of sampling bias. *Journal of Applied Psychology*, *76*, 229-238.
- Pettit, M. (2007). FMRI and BF meet FRE: Brain imaging and the Federal Rules of Evidence. *American Journal of Law & Medicine*, *33*(2-3), 319-340. doi:10.1177/009885880703300208.
- Raskin, D. C. (1989). Polygraph techniques for the detection of deception. *Psychological methods in criminal investigation and evidence* (pp. 247-296). New York, NY, US: Springer Publishing Co.

- Raskin, D. C., & Kircher, J. C. (2014). Validity of polygraph techniques and decision methods. *Credibility Assessment*, 63-129. doi:10.1016/b978-0-12-394433-7.00003-8.
- Reiman, J. H., & Leighton, P. (2013). *The rich get richer and the poor get prison: Ideology, class, and criminal justice* (10th ed.). New York, NY: Routledge/Taylor & Francis Group.
- Roberts, J. V., & Doob, A. N. (1990). News media influences on public views of sentencing. *Law and Human Behavior*, 14(5), 451-468. doi:10.1007/bf01044222.
- Rossi, P. H., Berk, R. A., & Campbell, A. (1997). Just punishments: Guideline sentences and normative consensus. *Journal of Quantitative Criminology*, 13(3), 267-290. doi:10.1007/bf02221093.
- Rovner, L. I. (1986). The accuracy of physiological detection of deception for subjects with prior knowledge. *Polygraph*, 15(1), 1-39.
- Rule 403, 28 U.S.C., § 403 (2018).
- Rule 702, 28 U.S.C., § 702 (2018).
- Sha Sha, R. & Fine, L.A. (2014). *Failed policies, forfeited futures: A nationwide scorecard on juvenile records*. Retrieved February 22, 2018, from <https://juvenilerecords.jlc.org/juvenilerecords/#!/map>.
- Silverthorne Lumber Company v. United States*. 251 U.S. 385.
- Staunton, C., & Hammond, S. (2011). An investigation of the guilty knowledge test polygraph examination. *Journal of Criminal Psychology*, 1(1), 1-14.
- Steinbrook, R. (1992). The polygraph test — A flawed diagnostic method. *New England Journal of Medicine*, 327(2), 122-123. doi:10.1056/nejm199207093270212
- The Oxford English dictionary*. (1989) (2nd ed.). Oxford: Clarendon Press.
- Tokunaga, H. (2016). *Fundamental statistics for the social and behavioral sciences*. Los Angeles: SAGE.

Tsoudis, Olga. (2000). Does majoring in criminal justice affect perceptions of criminal justice? *Journal of Criminal Justice Education*, 11(2). 225-236.

10.1080/10511250000084881.

United States Department of Commerce, Office of Audit and Evaluation. (2011). *Census 2010: Final report to Congress*. Washington, D.C.: U.S. Dept. of Commerce, Office of Inspector General, Office of Audit and Evaluation.

United States v. Scheffer. 523 US 303 (1998).

Varnson v. Satran, 368 N.W.2d 533 (1985).

Walsh, A. (2012). *Criminology: The essentials*. Los Angeles: SAGE.

What is EBDM? (2017). Retrieved April 12, 2017, from <https://info.nicic.gov/ebdm/>

Wisconsin v. Grignano and Loniello, 1935. Circuit Court of Columbia County.

Wood, J. L. (2009). Why public opinion of the criminal justice system is important.

In *Public Opinion and Criminal Justice* (Vol. 1).

Woolmington v. Director of Public Prosecutions, 1935, UKHL 462.

APPENDIX A

Statement of Informed Consent

Survey Participants Boise State University

You are being asked to take part in a survey regarding perceptions of the justice system, courts, polygraphs, and their credibility relative to other forms of evidence. The information collected from this survey will be used to determine the level of belief in the polygraph as well as perceptions on how it should be utilized. Questions will be asked about polygraph usage in criminal courts, usage outside criminal courts, allowance of examiner testimony, and comparison to other frequent types of evidence. This survey is being administered as a component of Mr. Jacob Schiess' thesis in the Department of Criminal Justice.

This survey represents a request for completely **voluntary participation** on your part; and your responses will remain totally confidential—neither your name nor any other identifying information will be asked or recorded nor will your identity be linked to any responses you provide. In addition, at any time during the survey, you may elect to skip any question asked or cease the entire process without penalty.

The Institutional Review Board at Boise State University, which reviews all proposed research involving human subjects, has approved this survey as meeting all requirements for the protection of the confidentiality of responses of survey participants.

If you have any questions concerning this survey, please feel free to contact us. Please write the following numbers down. The principal investigator, Mr. Jacob Schiess can be reached at (208) 890-4346. If you have any questions about your rights as a subject/participant in this research, or if you feel you have been placed at risk, you can contact the Office of Research Administration at (208) 426-5401. Dr. Lisa Bostaph, the co-investigator, can be reached at (208) 426-3886. If the survey places you under any measure of psychological stress for any reason, and you wish to seek counselling or other aid, University Health Services can be reached at (208) 426-1459. If you notice signs of stress in others, you can submit a report to the university CARE program at (208) 426-1527, or room 116 of the Norco building.

Before we begin, do you understand that your participation in its entirety or with regard to specific questions is completely voluntary, and that your responses will be kept confidential by the researchers?

Thank you in advance for your participation.

APPENDIX B

Survey Protocol

The following survey is confidential. Please do not add your student ID, first or last name, or any identifying characteristic beyond the demographic question in the survey. Please answer all questions truthfully and do not leave any blank.

For your convenience, the following definitions are available for your reference:

Polygraph: an instrument designed to detect and record changes in physiological characteristics, such as pulse or breathing rates

Control question test: a form of polygraph exam which compares a test subject's detected responses to relevant questions with their responses to control questions

Relevant question: an exam question which directly relates to the subject of investigation (e.g. "Did you rob John Smith?")

Control question: an exam question which does not relate to the subject of investigation, but often asks about generally illegal / undesirable acts (e.g. "Have you ever assaulted anyone in your life?")

Fingerprint evidence: any police-gathered and analyzed materials which require a professional fingerprint examiner to interpret

DNA evidence: all forms of blood work and body fluid analysis performed within forensic laboratories

Eyewitness evidence: all forms of individual testimony (either the victim or bystanders, excluding the distinct concept of expert witnesses)

Handwriting evidence: police investigation and subjective analysis of handwriting

Types of polygraph evidence:

- positive indications: passed tests (no deception detected)
- negative indications: failed tests (deception detected)
- lack of polygraph evidence (no test administered)

Survey Questions

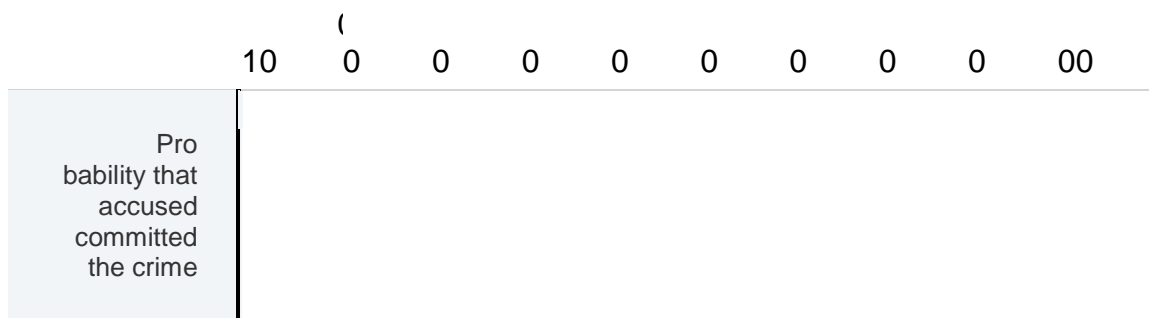
Q1

Based on the trial vignette provided, my verdict on the accused individual is?

- Guilty beyond a reasonable doubt
- Not guilty

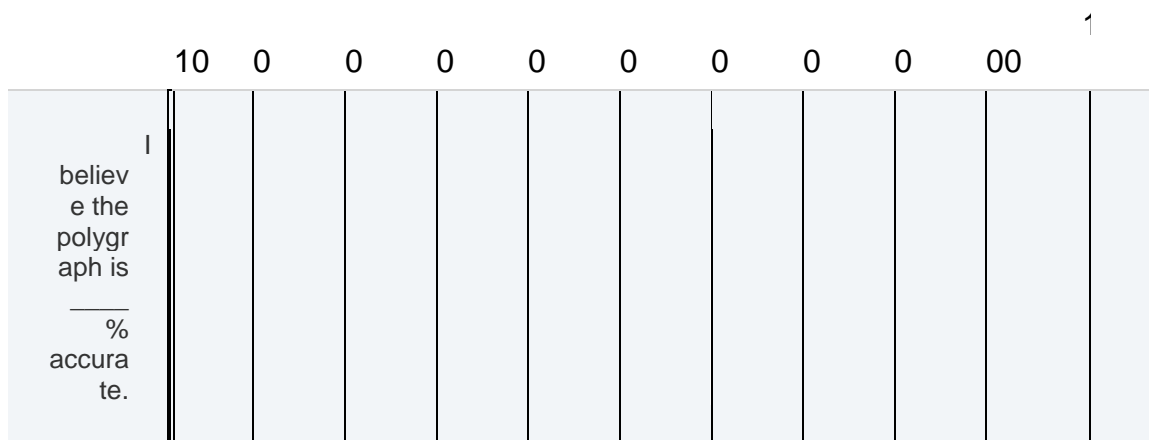
Q2

I believe there is a ____% probability (from 0 – 100) that the accused individual committed the crime.



Q3

On a range from 0 – 100%, how accurate do you believe the polygraph is? ____% (in general, not specific to this case)



Q4

How accurate (from 0 – 100%) do you believe you are at detecting deception informally? ____%

	10	0	0	0	0	0	0	0	0	00	1
I am _____% accurate at detecting lies informally.											

Q5

How much did the polygraph evidence in the vignette influence your decision on guilt?

- Not influenced at all
- Slightly influenced
- Moderately influenced
- Greatly influenced
- Extremely influenced

Q6

Should polygraph examiners be allowed in court to testify that the exam did not detect deception (i.e. the test subject was not lying)?

- Yes

- No

Q7

Should polygraph examiners be allowed in court to testify that the exam did detect deception (i.e. the test subject was lying)?

- Yes
- No

Q8

Which of the following statements best describes your opinion of the polygraph as evidence to be used in court?

- Sufficiently reliable to be the sole determinant (i.e. a polygraph exam result should be the only evidence needed to prove a case)
- Useful diagnostic tool when considered with other available information (i.e. an exam result can be helpful when combined with other evidence)
- Of questionable usefulness, entitled to little weight against other information (i.e. when present, an exam result should be viewed as weaker than most other evidence)
- Of no usefulness

Q9

In your opinion, is the polygraph more or less trustworthy than the other evidence types listed? Fill in the blanks in each sentence accordingly.

		Column Options	
		Polygraph evidence is ____	
		MORE trustworthy than	LESS trustworthy than
Fingerprint evidence		<input type="radio"/>	<input type="radio"/>
DNA evidence		<input type="radio"/>	<input type="radio"/>
Handwriting evidence		<input type="radio"/>	<input type="radio"/>
Eyewitness evidence		<input type="radio"/>	<input type="radio"/>

Q10

What is your age? If you would rather not disclose it, do not answer this question.

10 0 0 0 0 0 0 0 0 00

A											
ge in years											

Q11

What is your sex? If you would rather not disclose it, do not answer this question.

- Male
- Female

Q12

What level of course are you taking this survey as part of?

- 100-level
- 200-level
- 300-level
- 400-level

Q13

Was this survey administered to you in a criminal justice course? If you answer no, skip question 14.

- Yes
- No
- Not applicable

Q14

If this survey was administered to you in a criminal justice course, did it require upper-division standing?

- Yes
- No
- × Not applicable

Q15

Do you or a significant other (parent, spouse, immediate relative) work as a polygraph examiner or as a member of a criminal justice agency?

- Yes
- No
- Unsure / do not know

Q16

If you received an additional document which gave four examples of supportive / critical polygraph literature, how much did it affect your decision on guilt? If you did not receive said document, skip this question.

- Not influenced at all
- Slightly influenced
- Moderately influenced
- Greatly influenced

- Extremely influenced
- I did not receive the additional document

Q17

Do you have any additional comments regarding this study?

APPENDIX C

Trial Vignette

California resident John Doe, a citizen with no prior record, is being charged with the crime of first-degree homicide for killing his long-time girlfriend Jane Doe. State law allows the introduction of polygraph evidence to the jury and for jurors to draw their own conclusions from the results. There were no eyewitnesses to the crime, and the overall state of the evidence is relatively even. During the course of the trial, the defense brought forward evidence that John had taken and passed a respiration- and heartrate-based control question polygraph exam which examined his involvement in and knowledge of the crime. The polygraph examiner testified in court that, in their professional opinion, the results were inconclusive as to whether or not John had knowledge of the homicide. The examiner also testified that, in their professional opinion, John had been truthful when he denied having committed the homicide. The defense argued that John cannot be proven guilty beyond a reasonable doubt of this homicide and is pointing to the conclusively negative response from the polygraph (in regards to John's commission of the crime) as an indication of their client's innocence.

The prosecution's primary pieces of evidence were a bloodstained piece of fabric and a hair found on John Doe which was identified as belonging to the victim. The bloodstained fabric came from the shirt that Jane was wearing at the time of her death, and the blood was identified by DNA analysis to belong to her as well. The prosecution had also conclusively established that John had recently taken out a large life insurance policy on Jane. John was the last person to be seen with Jane prior to her death and was apprehended at a friend's house nearby within 24 hours of the police discovering Jane's body. The prosecution argued that John had the motive (collecting the insurance money),

and he killed Jane by strangling her to death with a rope. The police were not able to locate the rope in question, but the marks found on Jane's body make it very clear that she was strangled to death in that manner.

You are the foreman on a 12-person jury hearing this case. You and your fellow jurors are determining whether John Doe is guilty beyond a reasonable doubt. California law requires the verdict to be unanimous. The prosecution, in their closing arguments, attacked the defense's introduction of polygraph evidence as being improperly administered. They contend that the results can be ascribed to a 'friendly polygrapher', who they claim the defense located for the purposes of the trial. The defense argued in opposition that there was no such 'friendly polygrapher' effect. They point to the severity and notoriety of the crime as significant obstacles for any 'friendly polygrapher' to overcome.

APPENDIX D

Additional Literature Summary For Respondents

As with any topic in the justice system, there are critics and proponents to polygraphs and their usage in trials and other courtroom processes. When the instrument was first introduced to courts in 1923, it was prohibited in courtrooms under the Frye test, which required any evidentiary technology to have obtained “general acceptance” within the scientific community. Not until the case of *Daubert v. Merrell Dow Pharmaceuticals* (1993) was a new rule put into place which now requires judges to choose whether they will allow the polygraph in court, based on their own research and professional judgment. Literature exists that might lend credence to certain types of polygraph exams. Honts and Raskin (1988) found that 85% of innocent individuals and 92% of guilty individuals (in a mock crime scenario) were labelled as such by a control question test (CQT) polygraph. Furthermore, countermeasure attempts by test subjects to beat the polygraph were found to be ineffective. Myers et al. (2006) found that the concern some judges have regarding the polygraph’s supposedly extreme influence on jurors is also not present.

Problems exist for supporters of the polygraph. U.S. military courts and service branches prohibit the introduction of polygraph evidence as part 707 of their Rules of Evidence. The Supreme Court case of *U.S. v. Scheffer* (1998) resulted in a ruling that upheld the ban as constitutional. Within the civil / criminal sector, lawyers and proponents of the polygraph must demonstrate that evidence does not violate part 403 of the Federal Rules of Evidence (FRE). FRE 403 bans the introduction of evidence if its value as evidence is outweighed by the danger of introducing prejudice, if it misleads the jury, or confuses the issue at hand. Because polygraph examiners occasionally are required to testify in court on the exact meaning of test results, it is possible that judges could take exception with polygraphs under any of the three restrictions of FRE 403. The Supreme Court itself said in *U.S. v. Scheffer* (1998, pp. 4 - 9) that, “There is no consensus in the scientific community that the polygraph is reliable.”

APPENDIX E

This research was conducted under approval from the Institutional Review Board at
Boise State University, protocol #(044-SB17-149)