Author: Lamping, Julia L.

Title: Ethical Dilemmas and Obedience to Authority: Examining Evaluators’ Ethical Decision Making

The accompanying research report is submitted to the University of Wisconsin-Stout, Graduate School in partial completion of the requirements for the Graduate Degree/ Major: MS Applied Psychology Degree

Research Advisor: Tiffany Smith, Ph.D.

Submission Term/Year: Spring 2016

Number of Pages: 104


☑ I have adhered to the Graduate School Research Guide and have proofread my work. ☐ I understand that this research report must be officially approved by the Graduate School. Additionally, by signing and submitting this form, I (the author(s) or copyright owner) grant the University of Wisconsin-Stout the non-exclusive right to reproduce, translate, and/or distribute this submission (including abstract) worldwide in print and electronic format and in any medium, including but not limited to audio or video. If my research includes proprietary information, an agreement has been made between myself, the company, and the University to submit a thesis that meets course-specific learning outcomes and CAN be published. There will be no exceptions to this permission.

☑ I attest that the research report is my original work (that any copyrightable materials have been used with the permission of the original authors), and as such, it is automatically protected by the laws, rules, and regulations of the U.S. Copyright Office.

☑ My research advisor has approved the content and quality of this paper.

STUDENT:

NAME: Julia L Lamping DATE: 5/4/2016

ADVISOR: (Committee Chair if MS Plan A or EdS Thesis or Field Project/Problem):


This section for MS Plan A Thesis or EdS Thesis/Field Project papers only

Committee members (other than your advisor who is listed in the section above)

3. CMTE MEMBER’S NAME: DATE:

This section to be completed by the Graduate School

This final research report has been approved by the Graduate School.

Director, Office of Graduate Studies: DATE:
Lamping, Julia L. *Ethical Dilemmas and Obedience to Authority: Examining Evaluators’ Ethical Decision Making*

**Abstract**

Evaluators have been studied using self-report surveys documenting events of ethical dilemmas during projects (Morris, 2007; Morris & Clark, 2013), but little research has been done on the effects of unethical pressure from authority figures. Milgram’s (1974) work suggested that normal, every-day people will obey unethical requests proposed by someone who is seen as an authority figure. This study examined the effects of influential stakeholders making unethical requests on the decisions made by evaluators. Utilizing a situational judgement test, ethical dilemmas were presented to current American Evaluation Association members to examine ethical decision making. Two forms of the situational judgement test on ethics were developed; one where an authoritative figure is making an unethical request and one where the request is vague. Results provided partial evidence that obedience to authority and familiarity with the IRB process influences the decisions make in ethical dilemmas. Participants who received the Authority SJT-E form for Scenario Three where a client requests a break in participant confidentiality were over three times more likely to provide names of non-participants. Practitioners are recommended to utilize a standard method or strategy for making decisions in ethical dilemmas and be aware of client and stakeholder influence on situations.
Acknowledgments

Thank you to Tiffany Smith, who advised me through this process and kept me sane. Your interest in my thesis and research on evaluation pushed me forward when all I wanted to do was watch TV. Thank you for being not only my mentor and advisor, but for being a friend I could confide in. The reflective practice we engaged in was truly insightful and I learned a lot about myself and my competencies because of those meetings.

Thank you to Libby Smith, for your belief in me. I do not think I can successfully express how your words of encouragement pushed me forward. I appreciate everything you do for me. I especially appreciated the pictures of your two cats you shared; kittens are very motivating.

Thank you to Elizabeth Buchanan, for your encouragement and support throughout this process. Your expertise in ethics was insightful and well appreciated. I valued the outside opinion you provided. Thank you for your willingness to be on my committee and support me when I needed to get my Institutional Review Board proposal approved in the eleventh hour.

To Benjamin Weir and Ashley Hanke, thank you both for being my friends and my analysis buddies. You have gotten me through a lot these past two years. Thanks for always being on my team! I am so appreciative that I had you both by my side as I completed this adventure.

To my family and friends, thank you for the support and encouragement when I wanted to quit. Your warm thoughts pushed me forward. I hope I can make you as proud as you all make me.

Finally, a special thanks to Johnny Valles and our cat, Gatsby. Thank you. I love you.
# Table of Contents

Abstract ................................................................................................................................. 2  
List of Tables .......................................................................................................................... 8  
Chapter I: Introduction ......................................................................................................... 9  
  Statement of the Problem .................................................................................................... 9  
  Purpose of the Study .......................................................................................................... 9  
  Assumptions of the Study ................................................................................................. 9  
  Definition of Terms ......................................................................................................... 10  
  Limitations of the Study ................................................................................................. 10  
  Research Statements and Hypotheses ............................................................................ 10  
  Methodology .................................................................................................................... 11  
Chapter II: Literature Review ............................................................................................ 12  
  What is Evaluation? .......................................................................................................... 12  
  Ethics in Evaluation ......................................................................................................... 14  
  The Guiding Principles .................................................................................................... 15  
  The Essential Competencies for Program Evaluators .................................................... 17  
  The Program Evaluation Standards ................................................................................ 18  
  Evaluation Ethics Research ............................................................................................. 19  
  Obedience to Authority .................................................................................................... 23  
  The Current Study: Milgram and Evaluation .................................................................. 28  
  Hypotheses ..................................................................................................................... 30  
Chapter III: Methodology .................................................................................................. 31  
  Participants ....................................................................................................................... 31
Instrumentation .......................................................................................................................... 34
Situational Judgement Test-Ethics (SJT-E) ........................................................................ 34
Data Collection Procedure .................................................................................................... 36
Data Analysis .......................................................................................................................... 37

H1: Situations with Authoritative Pressure to Comply with Unethical Requests
(SJT-E: Authoritative) Will Illicit More Unethical Responses than
Situations without Pressure to Comply with Unethical Requests (SJT-E: Non-Authoritative) ........................................................................................................................................ 37

H2: Years as an Evaluator Practitioner will be Related to Amount of Unethical Responses .......................................................................................................................... 37

H3: Greater Familiarity with Ethical Codes, Standards, and Principles will
Influence how Ethical Participant's Responses to Scenarios will be ....... 37

Additional Analyses .............................................................................................................. 37

Summary .................................................................................................................................. 38

Chapter IV: Results .................................................................................................................. 39

Data Cleaning .......................................................................................................................... 39
Manipulation Check: Authoritative Presence and Ethically Problematic Ratings .......... 40
Results by Hypothesis ............................................................................................................. 43

H1: Situations with Authoritative Pressure to Comply with Unethical Requests
(SJT-E: Authoritative) will Illicit more Unethical Responses than
Situations without Pressure to Comply with Unethical Requests (SJT-E: Non-Authoritative) ........................................................................................................................................ 43
H2: Years as an Evaluator Practitioner will be Related to Amount of Unethical Responses .......................................................... 44

H3: Greater Familiarity with Ethical Codes, Standards, and Principles will Influence how Ethical Participants' Responses to Scenarios will be ...... 45

Rationale for Scenario Three Thematic Analysis ................................................................. 49

Summary ............................................................................................................................................. 55

Chapter V: Discussion, Conclusion and Recommendation ....................................................... 56

Discussion ........................................................................................................................................... 56

The IRB as an Authority in Decision Making ................................................................. 58

Obedience to Authority ........................................................................................................ 58

Limitations and Future Research ........................................................................................ 59

Implications for Practitioners .............................................................................................. 63

Figure 1: McDonald & Myrick Ethical Decision Making Table ............................................. 63

Figure 2: Novo Nordisk & University of Copenhagen Ethics Decision-Making Tool ....64

Recommendations for Education ......................................................................................... 64

Current Student Evaluators ................................................................................................. 64

Current Evaluation Practitioners ......................................................................................... 65

Conclusions ............................................................................................................................................. 65

References ......................................................................................................................................... 67

Appendix A: Implied Consent ................................................................................................. 71

Appendix B: SJT-E: Non-Authority ......................................................................................... 72

Appendix C: SJT-E: Authority ................................................................................................. 79

Appendix D: Debriefing ............................................................................................................. 86
Appendix E: Non-Authority SJT-E Form Rationale for Scenario One ................................................. 87
Appendix F: Authority SJT-E Form Rationale for Scenario One ..................................................... 90
Appendix G: Non-Authority SJT-E Form Rationale for Scenario Two .......................................... 93
Appendix H: Authority SJT-Form Rationale for Scenario Two ....................................................... 96
Appendix I: Non-Authority SJT-E Form Rationale for Scenario Four ........................................... 100
Appendix J: Authority SJT-E Form Rationale for Scenario Four ..................................................... 102
List of Tables

Table 1: Highest Degree Completed .............................................................................. 33
Table 2: Political Orientation ......................................................................................... 33
Table 3: Size of Current Organization or Employer .......................................................... 34
Table 4: Authority Presence Ratings by Scenario .............................................................. 41
Table 5: Ethically Problematic Ratings by Scenario .......................................................... 42
Table 6: Correlations between Authoritative Presence and Ethically Problematic Ratings by Scenarios .................................................................................................................. 42
Table 7: Scenario Three Responses by SJT-E Form .......................................................... 44
Table 8: Scenario One Responses and Familiarity with Ethical Codes, Standards, and Guidelines ......................................................................................................................... 47
Table 9: Scenario Two Responses and Familiarity with Ethical Codes, Standards, and Guidelines ......................................................................................................................... 47
Table 10: Scenario Three Responses and Familiarity with Ethical Codes, Standards, and Guidelines ......................................................................................................................... 48
Table 11: Scenario Four Responses and Familiarity with Ethical Codes, Standards, and Guidelines ......................................................................................................................... 48
Table 12: Scenario Three Rationale Themes: Non-Authority SJT-E .................................. 50
Table 13: Scenario Three Rationale Themes: Authority SJT-E ......................................... 52
Chapter I: Introduction

This study examines the influence authority presence has on decisions made in ethical dilemmas. Evaluators were surveyed using two forms of a situational judgement test on ethics. This chapter serves as an introduction to this study and provides a brief overview of the purpose of the study, terms used, and methodology.

Statement of the Problem

Stanley Milgram’s (1974) work showed that 65% of average citizens will obey an unethical request from an authority figure. While other professionals’ ethical decision making have been studied for obedience to authority influences (Blass & Schmitt, 2001; Burger, 2009; Dambrun & Vatiné, 2010), evaluators had yet to be researched. By understanding the influence of authoritative pressure on decisions made in ethically problematic situations, the field of evaluation can evolve to diminish the effects of client pressures.

Purpose of the Study

The purpose of this study is to determine the influence authority figures have on the ethical decision making of professional evaluators.

Assumptions of the Study

Evaluators work closely with clients to provided reports that are used to make data driven decisions about programming activities and funding. Due to the client-consultant relationship, ethically problematic situations are typical in business relationships. For this reason, it is important to study the influence the client-consultant relationship has in ethically problematic situations.
Definition of Terms

This study uses the theory of obedience to authority and literature regarding ethics in evaluation as its foundation. The following three terms are used in the study and are briefly described below.

Obedience to authority. Obedience to authority is compliance to requests of an authority figure (Milgram, 1974; Blass & Schmitt, 2001; Burger, 2009). Obeying these requests is not always voluntary in the sense that individuals have the free will to decline.

Evaluation. Evaluation is a process that gathers information to assess an evaluand’s merit, worth, and value (Alkin, 2012; Patton, 2012; Mertens, 2009).

Ethical dilemma. Ethical dilemmas are situations where moral behavior has wide applicability in a situation (Morris, 2011). Specifically, they are situations where an ethical issue has occurred and a decision needs to be made that has moral implications.

Limitations of the Study

The limitations of this study highlights the challenges when researching ethics in any field. In particular, using self-report as a means to study ethical decision making can open the study up to social desirability bias. Additionally, ethical standards and theories differ between and across cultures; making definitions of “right” and “wrong” contextually bound. Finally, comments regarding one of the manipulation check questions suggest participants did not fully understand the question or answered it assuming they were the authority in the scenarios.

Research Statements and Hypotheses

This study assessed the influence authoritative presence has on decisions made in ethical dilemmas. Specifically, this study assessed the effect authority figures have on compliance with unethical requests. Additionally, this study hypothesized a relationship between the sum of
unethical responses made on the situational judgement test scenarios and familiarity with ethical codes, standards, and guidelines, as well as years spent practicing evaluation. The three hypotheses that guided this study are:

- **H1**: Situations with authoritative pressure to comply with unethical requests (SJT-E: Authoritative) will illicit more unethical responses than situations without pressure to comply with unethical requests (SJT-E: Non-Authoritative).

- **H2**: Years as an evaluator practitioner will be related to amount of unethical responses.

- **H3**: Greater familiarity with ethical codes, standards, and principles will influence how ethical participant’s responses to scenarios will be.

**Methodology**

In order to examine the influence of clients on evaluators’ decision making in ethically problematic situations, four scenarios were developed using existing works as a base (Newman & Brown, 1996; Morris & Clark, 2013). Two versions on these four scenarios were created to compare similar scenarios between authoritative presence and a lack of authoritative presence. A sample of 2,000 current American Evaluation Association members were provided an online link either the authoritative situational judgement test or the non-authority situational judgement test. After three survey reminders, a final sample of 243 participant responses was usable for the study. Quantitative responses were cleaned and analyzed using a combination of Microsoft Access, Excel, and IBM SPSS. Qualitative comments were analyzed using NVivo software.
Chapter II: Literature Review

Milgram’s Obedience to Authority experiments (1974) resulted in an increased drive for studying the influence of authority figures on decision making, an increased awareness of ethical situations researchers face, and the approval of Institutional Review Board’s before collecting data. Milgram first set out to explain why Nazis in World War II followed their supervising officers’ orders to commit inhumane acts. What he found was shocking evidence that everyday people could commit horrible acts on another when pressured by someone they felt had authority over their behavior.

Since then, research has studied both partial replications of Milgram’s experiments (Blass & Schmitt, 2001; Burger, 2009) as well as studying the ethical decision making of professionals in applied settings (Newman & Brown, 1996; Morris & Clark, 2013; Brief, Dietz, Cohen, Pugh, & Vaslow, 2000; Peterson & Dietz, 2008). These replications have been both successful (Blass & Schmitt, 2001; Burger, 2009) and unsuccessful (Dambrun & Vatiné, 2010) at replicating Milgram’s findings. Researchers who have studied ethical decision making of professionals have found that the same authoritative obedience occurs even when codes and standards for ethical behaviors are present (Morris & Clark, 2013). Acknowledging the impact of authoritative pressure on decision making bias may be the first step to diminishing it. What evaluation is, obedience to authority, and ethical decision making and dilemmas of program evaluators will be addressed in the following sections.

What is Evaluation?

Evaluation is the systematic assessment of a program or project’s value, merit, worth, and significance (Mertens, 2009; Patton, 2012). Evaluation can sometimes be referred to as an inquiry process on the effectiveness of a program or product through use of applied research
methodology. Evaluations are useful in that they can answer a broad range of questions, depending on the focus. Such as: Is the evaluand effective? Was it implemented properly? Did it achieve its set goals and objectives? What are the outcomes from the program? What are the unexpected outcomes of the program? How had the participants of the program been affected? Were needs met? Can the program be improved? What worked and what didn’t? To answer the aforementioned questions, evaluators describe and assess programs and projects (Patton, 2012). Simply put, evaluations answer the “what, so what, now what?” questions about an evaluand (Patton, 2012, p. 3). “WHAT do we see?” “SO WHAT does it mean?” “NOW WHAT are our options for using this information?”

There are four important parts in the definition of evaluation. First, evaluation, like research, is a process that gathers information (Alkin, 2012). This includes using the appropriate methods and techniques to collect the information that is useful. Second, the information provided is used to make decisions, often resulting in action (rather than simply adding to theory or knowledge) (Alkin, 2012). Evaluators must provide valuable and useable recommendations based on the results of their evaluation (Alkin, 2012). Third, the information gathered and the recommended actions based on the results have to be presented to the final decision makers in a way that is usable and not misleading or confusing. The relaying of information to non-evaluators is an important step (Alkin, 2012). Simply throwing results of an evaluation, using technical jargon, into the lap of a client will not be useful for a stakeholder. Likewise, data provided should be visualized in a way that is easy to understand and correct interpretations are required. Finally, evaluations differ based on the information needed to be gathered and the type of decisions available to be made (Alkin, 2012). An evaluation of a prison rehabilitation program will differ greatly from an evaluation of a school districts mentoring program. Therefore, there
are multiple approaches to evaluations (Alkin, 2012). A goal or outcome of one program may even differ from a similarly structured program elsewhere. The appropriate approach and methodology used depends on the situation as well as the evaluator (Alkin, 2012). There are five general types of evaluation: “studies of program outcomes; cost-effectiveness studies and cost-benefit analysis; implementation studies; consumer satisfaction studies; and needs studies” (Clarke & Dawson, 1999). Depending on the evaluand, the stakeholders, and the situation itself, the type of evaluation can change.

Evaluation is also a form of applied social research, focusing on studying effectiveness using existing knowledge and theory to inform and guide action (Clarke & Dawson, 1999). A line needs to be drawn between generic research and evaluation though. Generic, or basic, research adds to field’s existing theories and knowledge. Evaluation research uses existing theory and knowledge, applied research methodology, and uses data to make informed decisions (Clarke & Dawson, 1999). Simply, evaluation research typically does not add to its own field’s theories. Program evaluations are situationally dependent on the context of the organizational and political factors in which they are conducted, unlike research (Mertens, 2009). However, ethics in research with human subjects applies to both evaluations and research. Basic psychological research requires review of the Institutional Review Board (IRB) and compliance with APA Ethical Principles and Code of Conduct. Likewise, evaluation research has its own specific set of codes, guidelines, and principles for dealing with ethics and compliance.

**Ethics in Evaluation**

Evaluation research is highly customizable to the program being evaluated, thus, evaluators must be well versed in the ethical norms of the organization and participant rights while designing program evaluations for clients. Morris’ (Morris & Jacobs, 2000; Morris, 2007,
many years of research, on the ethical dilemmas evaluation researchers face, has added significantly to the field of evaluation research. The current paper adds to the body of research on evaluators and ethical decision making by linking it to the pressure to obey unethical requests by high-stakes clients (obedience to authority).

Ethics is usually defined in terms of “wrong” or “immoral” behaviors or decisions (Newman & Brown, 1996); much like a “no-no” list of actions for professionals to avoid. An ethical dilemma is then a situation in which a choice must be made between at least two possible decisions that are equally unpleasant and blur the lines of morality (Newman & Brown, 1996; Morris, 2007). Ethical decisions are the “morally good” or “the right thing to do” and can be easier to spot when the subject of the dilemma is someone else (Morris, 2007).

Professional evaluators are held to the ethical and professional standards of: The AEA’s Guiding Principles (AEA, 2004), the Essential Competencies for Program Evaluators (Stevahn, King, Ghere, & Minnema, 2005), and the Program Evaluation Standards (Yarbrough, Shulha, Hopson, & Caruthers, 2010). The following sections break down each of these ethical guidelines in detail.

**The Guiding Principles**

The Guiding Principles were put together by Scheirer, Newman, Shadish, and Wye (1995) from the AEA after analyzing other professional societies’ standards (Scheirer et al., 1995). These Principles were officially recognized by AEA in 1994 after a year of drafting the Guiding Principles. The Guiding Principles are used to guide professionals in everyday evaluations and reactions when an ethical dilemma occurs (Scheirer et al., 1995). The five principles are Systematic Inquiry, Competence, Integrity and Honesty, Respect for People, and Responsibilities for General and Public Welfare (AEA, 2004; Scheirer et al., 1995).
Systematic Inquiry means evaluators should use appropriate methodology and data analysis when inquiring around the evaluand (Scheirer et al., 1995). The highest standards for methodology should be used and limitations of those methodologies should be discussed with the client and in detail in the evaluation report to allow others to “understand, interpret, and critique their work” (AEA, 2004, p. 2). Competence refers to the competent performance of the evaluator. Evaluators must possess the necessary knowledge, skills, and abilities to complete the evaluation project appropriately as well as demonstrate cultural competence for working in different groups. Evaluators without the necessary skills should decline work tasks that are outside their abilities but strive to maintain and improve their competencies (AEA, 2004, p. 2).

Integrity and Honesty refers to how evaluators should be honest and uphold the integrity of the evaluation throughout the entire process (Scheirer et al., 1995). This includes creating and facilitating open and honest communication with clients and stakeholders throughout the evaluation about all aspects of the project (AEA, 2004, p. 2). The Guiding Principle, Integrity and Honesty, also addresses resolving conflict and concerns in evaluations and clearing up misleading information. Respect for People refers to respecting all individuals they interact with. Evaluators should “respect the security, dignity, and self-worth of the respondent, program participants, clients, and other stakeholders” (Scheirer et al., 1995, p. 2). Finally, the Responsibilities for General and Public Welfare mean that evaluators are responsible for articulating and being concerned with the diversity of interests and values related to the general and public (Scheirer et al., 1995). Simply put, according to the Guiding Principles, evaluators should do the most good for the most people.
The Essential Competencies for Program Evaluators

The Essential Competencies for Program Evaluators were developed by Stevahn, King, Ghere, and Minnema (2005) starting in the early 2000s. The systematic analysis of various meetings, sessions, and comments resulted in the initial taxonomy of the Evaluation Competencies. The current competencies were created so that they would be user-friendly, cross-reference the evaluation standards endorsed by evaluation associations, support existing evaluation association standards, are precise, and can be disseminated in a way that is both timely and open to continuous improvement (Stevahn, et al., 2005).

The Essential Competencies are split into six competencies; (a) professional practice, (b) systematic inquiry, (c) situational analysis, (d) project management, (e) reflective practice, and (f) interpersonal competence (Stevahn, King, Ghere, & Minnema, 2005). Professional practice says evaluators should adhere to evaluation standards and ethics (Stevahn, King, Ghere, & Minnema, 2005). Systematic inquiry refers to the technical side of evaluation research- the methodology and analysis (Stevahn, King, Ghere, & Minnema, 2005). Situational analysis means the context of the situation should be attended to so an evaluator doesn’t miss an interest or issue in a particular evaluation setting. Project management refers to conducting an evaluation. Specifically, the budgeting, coordination, and supervision of the evaluation process are vitally important (Stevahn, King, Ghere, & Minnema, 2005). Reflective practice says evaluators should focus on their awareness of the evaluation and their areas of growth. Reflective practice can be used to engage in professional development to strengthen an evaluator’s areas of weakness (Stevahn, King, Ghere, & Minnema, 2005). Interpersonal competence refers to the needed people skills evaluators require when conducting an evaluation. These skills include
communication, negotiation, conflict and collaboration facilitation, and cross-cultural skills (Stevahn, King, Ghere, & Minnema, 2005).

The Program Evaluation Standards

The Joint Committee on Standards for Educational Evaluation (JCSEE) was formed in 1974, now represented and supported by 17 sponsoring organizations, to develop and implement evaluation standards (Sanders, 1994). This committee reviewed the perspectives of a diverse group of stakeholders from inside and outside North America. Every ten years, the standards go through another review process in order to maintain American National Standards Institute approval (Sanders, 1994). The result of the JCSEE is the Program Evaluation Standards. In particular, the Program Evaluation Standards are accredited to Yarbrough, Shulha, Hopson, and Caruthers (2011) acting in a joint committee on standards for educational evaluation.

The Program Evaluation Standards are split into five categories: Utility, Feasibility, Propriety, Accuracy, and Evaluation Accountability. The Utility Standards focus on the usability of evaluation products for program stakeholders. Specifically, these standards include the evaluator’s credibility, attention, negotiation, values, relevant information, process, communication and reporting, and their concern for consequences (Yarbrough, et al., 2011). Feasibility Standards focus around the effectiveness and efficiency of the evaluation. Propriety Standards refer to how evaluators should be proper, fair, and respectful to the stakeholders and protect human rights throughout the evaluation process (Yarbrough, et al., 2011). Any conflicts of interest must be honestly discussed and all resources must be accounted for.

Accuracy Standards focus on the validity and reliability of the results of the evaluation (Yarbrough, et al., 2011). Interpretations and judgements about quality should be made through appropriate research design and analysis. Finally, Evaluation Accountability Standards refer to
the appropriate documentation and meta-evaluation process to hold evaluators accountable for their processes and evaluation products (Yarbrough, et al., 2011). Even experienced evaluators have faced ethical dilemmas that challenge or threaten one or more of the above principles and standards (Newman & Brown, 1996; Morris, 2007).

**Evaluation Ethics Research**

Program evaluators are professionals working to assess the merit, worth, value and/or significance of a program (Newman & Brown, 1996; Morris, 2007). Evaluators are, at best, trained in ethics by a “chapter or two in a test or a semester session or two in a professional course” (Newman & Brown, 1996, p. 3). While any exposure to ethics and ethical decision making is better than no training at all, program evaluators often enter the field without the knowledge, skills, or abilities to make ethical decisions when faced with issues (Mabry, 1999; Newman & Brown, 1996). Mabry (1999) suggested ethical dilemmas are resolved from within, although standards, principles, codes, guidelines, and collective reflection all help.

Morris and Jacobs (2000) examined the factors that influence how evaluators interpret ethical dilemmas and had participants give explanations for their decisions which were analyzed using the Guiding Principles for Evaluators as the major themes. Three research situations were given to the U.S. members of AEA in 1999 and 391 responses were usable for analysis. The research situations included scenarios where the evaluator alters a section of the final report after pressure from a stakeholder, the evaluator does not involve stakeholder participation in the evaluation process, and where the evaluator obtains passive consent (not active consent) when evaluating a school-based youth program (Morris & Jacobs, 2000). Participants rated on a scale from 1 (definitely problematic) to 5 (definitely not problematic) each situation’s potential to be ethically problematic. After each Likert item, participants were given the opportunity to provide
open ended responses as to why they chose their answer. Participants also rated how familiar they were with the AEA’s guiding principles and their political orientation. Additional information obtained included years worked in evaluation and how many evaluations they have conducted (Morris & Jacobs, 2000). Morris and Jacobs (2000) found that 69% of participants rated the evaluator in the passive consent situation as definitely or probably ethically problematic. A little over 50% rated the evaluator that revised their final report under pressure as problematic, and then only 39% rated the evaluator who did not involve their stakeholders as ethically problematic.

The Guiding Principles for Evaluators were used to organize and analysis into specific themes when comparing participant’s responses to the open ended explanations (Morris & Jacobs, 2000). Morris and Jacobs (2000) found that many guiding principles could be used as arguments for either side (definitely problematic or definitely not problematic). For example, 45% of participants who believed the passive consent situation violated the guiding principle Respect for People, while 43% of those who believed the same situation were not an ethical problem felt the guiding principle Respect for People was the explanation. Specifically, the 45% felt the evaluator violated informed consent despite his/her knowledge of its limitations, while the 43% felt passive consent is ethical when obtaining consent (Morris & Jacobs, 2000).

The scenario in Morris and Jacob’s 2000 study found that the about half of participants felt the revising a report was ethically problematic. In order to delve deeper into report (mis)handlings, Morris and Clark (2013) studied the pressure to misrepresent findings evaluators encountered. Morris and Clark (2013) surveyed AEA members on their experience with ethical dilemmas. Misrepresentation of findings goes against the AEA Guiding Principles (AEA, 2004) as well as the Program Evaluation Standards (Yarbrough, Shulha, Hopson, & Caruthers, 2011).
Morris and Clark (2013) found that 42% of participants reported being pressured to misrepresent evaluation findings. When asked about solutions to the pressure, half of the participants felt “nothing could have been done” while 30% replied preventative measures could have been taken, and 20% were unsure (Morris & Clark, 2013). Creating a detailed protocol in the formal contract that specifies how to deal with pressure to misrepresent and revisions on final reports was a suggestion for preventing situations where misrepresentation pressure occurs (Morris & Clark, 2013).

Morris and Clark (2013) also found that the pressure to misrepresent findings was more likely to come from clients of the project (69%) compared to pressure from the director of the program (16%) or others such as staff members, evaluation supervisors or leaders, coworkers, or other stakeholders (15%). There is high pressure to please clients, particularly when stakes for delivering “bad news” are high for a well-liked program (Morris & Clark, 2013). Morris and Clark (2013) briefly mentioned the pressure from clients being particularly forceful because clients are paying for services. Morris and Clark (2013) found that of the reported misrepresentation cases studied, 15% of participants said misrepresentation occurred while 3% were fired from the evaluation.

Consequences for completing an evaluation project where the client is unhappy or dissatisfied with the results or dissemination can include outcomes from distrust for evaluators in general to refusing to hire evaluators from that organization again (McDonald & Myrick, 2008). On the other hand, complying with unethical requests, like misrepresentation of findings, can leave the client satisfied with the final report and more likely to use recommendations or evaluation services in the future (McDonald & Myrick, 2008). The dynamic between keeping a
friend and making an enemy out of an existing client is seen in McDonald and Myrick (2008) through the lens of a professor and a team of graduate students on their first evaluation project.

McDonald and Myrick (2008) also stressed stakeholder analysis to prevent and solve ethical dilemmas and concerns. The case study highlighted the events of a student evaluator group’s first project. The student evaluators were working on their first real-client project; an assessment of a diversity training program on campus. During data collection, a comment was found regarding an issue with one of the diversity speakers in the program. The comment said that one diversity speaker was making racial comments and perpetuating stereotypes during lectures (McDonald & Myrick, 2008). When reporting these results to their client, the head of the diversity program on campus, and to their mentoring professor, the client demanded the speaker’s identity. The mentoring professor suddenly found himself, and his students, in an ethical dilemma. On the one hand, the professor was close friends with the client and understood the concerns and ramifications of leaving the speaker to continue lecturing on diversity. On the other, the diversity speaker was a participant in the study and was promised confidentiality. The choice between angering their client and breaking confidentiality was difficult for the mentoring professor and his student evaluators (McDonald & Myrick, 2008). McDonald and Myrick (2008) highlighted the importance of keeping to the ethical and professional principles and standards for evaluators, and creating clear protocols for situations that may arise during projects.

At times, the ethical or “right” decision can be hazy or unclear, or the pressure by others to make a certain decision can be great (Newman & Brown, 1996). It is important to study the effects of authoritative pressure on evaluators. One would hope an ethical choice is always made, and that the principles and standards for evaluators adhered to. However, lack of traditional
ethical training and the unfamiliarity of the situation may make it easier for clients or project
managers to pressure evaluators to make immoral decisions.

**Obedience to Authority**

Obedience to authority refers to complying with requests of another who is perceived to have an authoritative stance (Milgram, 1974; Blass & Schmitt, 2001; Burger, 2009). Obeying these requests are voluntary in the sense that individuals have the free will to decline; however, the added presence of the authority figure elicits compliance regardless of whether the individual wants to decline. Milgram believed people were socialized at a young age to obey authority figures without question (Milgram, 1963). At the time, children were taught not to question adults or those with a higher social or political rank (Milgram, 1963), so obeying a legitimate authority figure may have felt rewarding.

Milgram’s (1974) Obedience to Authority experiments at Yale resulted in shocking insights into the moral conduct of everyday people and the requirement to protect participants in research. Participants in Milgram’s obedience to authority experiments were deceived into thinking they were teaching another participant to memorize a list of works and its pairs. When the learner in the other room, attached to the fake-shock generator, answered with an incorrect word to its pair, the participant administered a shock of electricity into their hand. As the wrong answers continued, the participant was instructed to increase the level of shocks. The shock generator ranged from 15 volts (Slight Shock), to 375 volts (Danger: Severe Shock), to 450 volts (XXX) (Milgram, 1974). The actual learner in the experiments was a confederate to the experiment. However, the learner’s protests to being involved in the experiment increased as the volts administered increased. When the participants of the experiments attempted to refuse to shock another person, the experimenter, clad in a lab coat, responded with one of four prods to
continue the shock experiment. Milgram found that normal individuals will continue to administer shocks when instructed to by a legitimate looking individual (the experimenter). Of the 40 participants in Milgram’s experiment, all of them continued past the expected breakoff point (275 volts), five stopped after administering the 300 volt level, four went one level further, two went another level further. A total of 14 of the 40 participants stopped the experiment before reaching the end of the voltage levels (450 volts) and therefore were counted as defiant (Milgram, 1963). Sixty-five percent of participants in the study were counted as fully compliant to the experimenter (Milgram, 1963).

Many studies regarding obedience to authority focus on partial replications of Milgram’s experiments (Blass & Schmitt, 2001; Burger, 2009; Dambrun & Vatiné, 2010). Specifically, the studies included situations where the participant is told they are involved in a study on learning and are instructed to use a virtual shock generator under the supervision of an experimenter (Blass & Schmitt, 2001; Burger, 2009; Dambrun & Vatiné, 2010). The power of the situation is greatly applicable to everyday situations and should be examined in the context of professional decision making. For example, Burger (2009) partially replicated Milgram’s (1974) Yale experiments utilizing a virtual game. The replication Burger (2009) designed tested Milgram’s obedience to authority results against his own while also adding in females to the participant pool. Burger (2009) found participants in his study obeyed at the same proportion as Milgram found decades before.

Burger (2009) concluded that even though he did not have a 450 volt option as in Milgram’s experiments, the individuals who obeyed until the end of his experiments likely hit a “point of no return” at 150 volts. That is to say, the virtual experimenter (the authority figure) had used the “foot-in-the-door” method to gain compliance. Participants felt obligated to
continue after hitting a metaphorical “point of no return” because they had complied with the increasing requests for higher volts already. The obedience study demonstrated how most individuals underestimate the power of the situation on their behavior (Burger, 2009).

In Burger’s (2009) replication, gender (male/female) and personality (empathy and need for control) differences were taken into account; unlike in Milgram’s original studies. A total of 29 men and 41 women participated in Burgers (2009) replication. Burger’s experiment stopped at the “point of no return” (150 volts); participants who stopped prior to 150 volts were considered non-compliant (deviant) while those who continued past 150 volts were assumed to have passed a “point of no return.” Of the 70 participants, 70% continued past 150 volts and were stopped by the experimenter prior to the next level. The rate of compliance (70%) did not differ significantly from Milgram’s own compliance rate for the 150 volt level (82.5%) (Burger, 2009). When comparing personality of those who complied with those that were defiant, Burger (2009) found participants did not differ on their empathetic concerns score, or their desire for control score. Additionally, Burger (2009) found no gender differences on how far participants were willing to go before stopping or being stopped.

Another Milgram-inspired replication was done by Blass and Schmitt in 2001; studying participants explanations for obedience in Milgram’s study. Blass and Schmitt (2001) randomly assigned 38 undergraduate students in an experimental psychology course to one of two conditions: Choice or Likelihood. All the participants watched an edited version of the Milgram documentary, Obedience, which included the beginning of the experiment (the explanation and set up) up until the participant administered the 180 voltage in the film. “Choice” group participants were tasked with putting a check mark next to one of six explanations that best describe why the participant in the film complied with the experimenter’s request. Participants in
the “likelihood” group were also given the same six explanations but had to rate each on a Likert-scale from 1 (extremely unlikely) to 5 (extremely likely) on the likelihood the participant in the film complied for that reason. Blass and Schmitt (2001) found that both groups of participants in their experiment felt the legitimacy of the experimenters explained the participant’s compliance more so than the other five explanations.

A gamified partial replication of Milgram’s (1974) experiment was done by Dambrun and Vatiné (2010). Dambrun and Vatiné (2010) studied 31 undergraduate students in France regarding their obedience to a virtual authority figure in an online version of Milgram’s experiments and the influence of race on their obedience rates. These students were given the same cover story and instructions as Milgram’s participants, however, they were told explicitly that the learner in the online game was an actor and was in no actual pain or harm throughout the study. Dambrun and Vatiné (2010) also manipulated the ethnicity of the victim (French victim vs. North African victim) and the visibility of the victim (visible but heard vs. visible and heard). The participants were more likely to fully obey instructions to shock virtual participants when the participant wasn’t seen and only heard. There were no significant differences or interaction effect when the victim’s ethnicity changed from French to North African. When comparing their hidden-victim experiments to Milgram’s (1974) vocal feedback only (also hidden-victim) experiment, the results did not differ significantly. Dambrun and Vatiné (2010) found that 53% of their participants obeyed fully when the victim was hidden and 62.5% of Milgram’s participants fully obeyed in his vocal feedback experiments.

In the visible-victim experiments, only 13% of participants fully obeyed and administered the highest voltage; significantly different from Milgram’s 40% in his experiments. Dambrun and Vatiné (2010) believe the difference in obedience rates were likely due to the victim’s
extreme reactions compared to Milgram’s visible victims. Student’s in Dambrun and Vatiné’s (2010) experiment acted more violently in response to the shocks compared to Milgram’s.

Another interesting finding was that the higher the shocks participants administered to the virtual victim, the more the participant attributed responsibility of that action to the experimenter and the victim. Additionally, Dambrun and Vatiné (2010) found that state-anger and right-wing authoritarians were both significant predictors of obeying to the final voltage of the experiment.

Ties to the “real-world” between obedience to authority and organizational research can be seen in Brief, Dietz, Cohen, Pugh, and Vaslow’s research (2000). Participants in Brief and colleagues’ study (2000) were told to recommend applicants for hiring based on resumes and a letter from the CEO instructing them, indirectly, to avoid hiring any Black applicants. Brief and colleagues (2000) found that participants viewed the CEO as a legitimate authority figure, discriminated against Black applicants, and claimed the authority figure’s racist request was justification for their actions. Specifically, the CEO (authority figure in the study) instructed participants to keep the work environment harmonious and that organization members and customers felt more comfortable dealing with others of their own race. Brief and colleagues (2000) tied the obedience to authority literature to practical events. Specifically, their participants obeyed an unethical, and illegal, request from someone they viewed as an authority figure. Interestingly, participants did not obey the request to racially discriminate against Blacks when the request came from an illegitimate figure (a current employee) (Brief, et al., 2000). This suggests legitimate, authoritative individuals who are perceived to have power in the situation are easier to obey.

hypothesized that participants who receive direction from an authority figure in their organization to discriminate based on demographic criteria will comply by selecting fewer applicants with the non-preferred demographics. Peterson and Dietz (2008) also looked at the influence of organizational commitment on compliance to authority figures. Organizational commitment is the emotional bond an employee has to their organization (Peterson & Dietz, 2008).

Peterson and Dietz (2008) predicted that the more committed to an organization a participant is, the more likely they will comply with directions to discriminate based on demographic criteria. One hundred and nine participants received the same assignment to pick applicants out based on qualifications. Half of the participants were given directions by a principal that implied to select applicants from East Germany over West Germany. Results suggested that applicant discrimination was found in higher rates from participants who received the direction from the principal and had high organizational commitment compared to those with low commitment and those that did not receive the directions (Peterson & Dietz, 2008).

The Current Study: Milgram and Evaluation

Evaluation researchers can learn from the long past of social psychologists regarding social influences and persuasion (Mark, Donaldson, & Campbell, 2011). Likewise, social psychologists can learn from the social influence and persuasion that occurs in organizations outside labs (Mark, Donaldson, & Campbell, 2011). The power of fundamental social theories should not be underestimated. Previous research has studied peer pressure, organizational culture and ethical environments, as well as personal morals and confidence. Few have studied the influence of authority figures on decisions made during ethical dilemmas in evaluation. For this reason, it is important to look at the influence of authoritative pressure on evaluator’s
decision making and solutions to ethical dilemmas. Practitioners are susceptible to obeying requests and falling to pressure like any other profession. However, evaluation practitioners are providing results-based recommendations and consultation to heads of programs and should be held accountable to the standards and principles for maintaining integrity. The present study tests the ethical decision making of evaluators with and without the added pressure of authorities, high stakes stakeholders. Recommendations and suggestions were made based on the present study’s results.

A situational judgement test using ethically problematic scenarios was utilized to examine ethical decision making of evaluators. These scenarios were developed using previous scenarios as inspiration or models and reviewed by subject matter experts prior to piloting. Response choices for each scenario were developed through subject matter expert collaboration. Each situation has two possible responses to choose from. Participants will be instructed to choose the response that best describes what they would do in each scenario.

Two forms of the scenarios was created. The first, Situation Judgement Test- Ethics: Non- Authoritative, has no added legitimate authority presence making an unethical request. Rather, the scenarios vaguely describe the ethical dilemma without additive pressure from an authority figure. The second form, Situational Judgement Test- Ethics: Authoritative, uses the same scenarios with the additive pressure from a legitimate authority figure to make an unethical decision. Each scenario will ask participants to explain their reasoning behind their response.

In addition, each scenario has a question that asks the participant to rate how ethically problematic they feel each situation was (1 = ethically problematic, 5 = not ethically problematic). Participants also rated how much of an authoritative presence they felt each scenario has (1 = no authoritative presence, 5 = a lot of authoritative presence). Demographic
information will be obtained concerning the participants sex, years in practice, current employment status, how familiar they are with current ethical codes, standards, and principles, and if they identify as primarily an internal or external evaluator.

**Hypotheses**

This study strove to provide evidence of authority figures influence on decisions made in ethical dilemmas, specifically, compliance with unethical requests. This study also assessed the relationship between decisions made in each scenario and the amount of time spent practicing evaluation in years. Additionally, this study assessed the relationship between familiarity with ethical codes, standards, and principles and the sum of unethical responses made on the situational judgement test scenarios. The three hypotheses that guided this study are:

- **H1**: Situations with authoritative pressure to comply with unethical requests (SJT-E: Authoritative) will illicit more unethical responses than situations without pressure to comply with unethical requests (SJT-E: Non-Authoritative).

- **H2**: Years as an evaluator practitioner will be related to amount of unethical responses.

- **H3**: Greater familiarity with ethical codes, standards, and principles will influence how ethical participant’s responses to scenarios will be.
Chapter III: Methodology

To examine the effects authority figures and familiarity with ethical codes and standards have on evaluator decision making in ethically problematic situations, a random sample of 2,000 American Evaluation Association (AEA) members were invited to participate in a situational judgement test on ethics (SJT-E) and a total of 243 provided complete responses. The SJT-E utilized four ethically problematic scenarios which had participants choose between an ethical and unethical response and then provide rationale for their chosen option. This section describes the participants, the SJT-E forms, data collection procedure, and the data analysis procedure.

Participants

Participants were comprised of current program evaluators who are non-student members of the AEA. The AEA is a professional association of evaluators to share best practices, professionally develop, and network. AEA evaluators work in the evaluation research field conducting program evaluations and research on evaluation best practice. There are approximately 7,000 members of AEA in the United States and in over 60 foreign countries (About AEA, 2015, Eval.org).

A random sample of 2,000 AEA members received an email invitation to participate in the study. The link to the testing survey site will emphasize that participating in the proposed study will be confidential. Reminder emails were sent out one and three weeks after the initial email invitation. Morris and Clark (2013) also used AEA members as participants and obtained an overall response rate of 37%, which suggested this method of sampling evaluators would be sufficient. Stout Institutional Review Board and AEA Institutional Review Board approval were obtained before participants were contacted for data collection. A total of 243 members fully participated in the study; a response rate of 12.15% (or 243/2000).
A total of 164 participants opened the Non-authority situational judgement test. Of those 164, 12 respondents were under the age of eighteen and were taken to the end of the online survey. Another 29 did not complete any of the test questions or only responded to the first and were removed during cleaning. Additionally, seven more were removed for answering only the first two scenarios and then leaving the rest of the SJT-E blank. A total of 115 participants provided usable responses to the Non-Authority form (final response rate of 11.50% or 115/1,000 possible).

A total of 178 participants opened the Authority situational judgement test. Of those 178, 12 responded being under 18 years of age and were taken to the end of the online survey. Another 36 were removed for not completing over 90% of the questions. Additionally, three did not provide responses to quantitative questions and used the qualitative questions as means to criticize the scenario choices or refuse to choose either option; these participants were excluded from data analysis. A total of 128 participants provided usable responses to the Authority form (final response rate of 12.80% or 128/1000).

One hundred and seventy five participants identify their sex as female (73.2%), sixty identified as male (25.1%), and one identified as intersex (.4%) while three did not respond to the question. With regards to gender identity, 175 (73.2%) identified as woman, while 58 identified as men (24.3%), two preferred not to disclose their gender (.8%), and four did not respond to the question. The majority of participants have earned a Ph.D. or equivalent (n = 122, 51%; see Table 1 below). Ninety-nine participants have earned a Masters’ degree (41.4%) while five have earned a bachelor’s degree (2.1%). The majority of participants also identified as liberal or very liberal (38.1% and 28.5% respectively; see Table 2 below.)
One hundred and eighty four (77%) of the participants were employed for an organization while 37 were self-employed or independent consultants, nine were students, three were retired, and one was unemployed. The size of the organizations participants’ currently work for varied, though the majority of participants work at an organization with at least 100 employees (see Table 3). Evaluation was a primary job task for the majority of participants (82.8%) and the majority considered themselves an external evaluator (51%), whereas 41% considered themselves an internal evaluator and 6.7% responded with “other” (participants responded that they were both internal and external evaluators). The average amount of years as an evaluation practitioner was 12.58 years \((SD = 9.64)\) and ranged from 0 to 50 years.

Table 1

*Highest Degree Completed*

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bachelors</td>
<td>5</td>
<td>2.1</td>
</tr>
<tr>
<td>Masters</td>
<td>99</td>
<td>41.4</td>
</tr>
<tr>
<td>Ph.D. or Equivalent</td>
<td>122</td>
<td>51.0</td>
</tr>
<tr>
<td>Other</td>
<td>10</td>
<td>4.5</td>
</tr>
</tbody>
</table>

Table 2

*Political Orientation*

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Conservative</td>
<td>1</td>
<td>.4</td>
</tr>
<tr>
<td>Conservative</td>
<td>10</td>
<td>4.2</td>
</tr>
<tr>
<td>Moderate</td>
<td>61</td>
<td>25.5</td>
</tr>
<tr>
<td>Liberal</td>
<td>91</td>
<td>38.1</td>
</tr>
<tr>
<td>Very Liberal</td>
<td>68</td>
<td>28.7</td>
</tr>
</tbody>
</table>
Table 3

*Size of Current Organization or Employer*

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 or self-employed</td>
<td>34</td>
<td>14.2</td>
</tr>
<tr>
<td>2-9</td>
<td>24</td>
<td>10.0</td>
</tr>
<tr>
<td>10-24</td>
<td>19</td>
<td>7.9</td>
</tr>
<tr>
<td>25-99</td>
<td>22</td>
<td>9.2</td>
</tr>
<tr>
<td>100-499</td>
<td>30</td>
<td>12.6</td>
</tr>
<tr>
<td>500-999</td>
<td>16</td>
<td>6.7</td>
</tr>
<tr>
<td>1000-4999</td>
<td>43</td>
<td>18.0</td>
</tr>
<tr>
<td>5000+</td>
<td>44</td>
<td>18.4</td>
</tr>
</tbody>
</table>

**Instrumentation**

This study measured compliance to authoritative requests and presence using a situational judgement test. In addition, questions regarding ethical codes, standards, and principles familiarity, demographic questions, and manipulation check questions were included after the scenarios. The following paragraphs describe the instrumentation used in this study.

**Situational judgement test-ethics (SJT-E).** A situational judgement test regarding ethical dilemmas was created for the purpose of the study. Situational judgement tests are videos or written scenarios to which participants respond. In the present study, the situations were presented via written story. There were four total scenarios for which participants chose a behavioral response that best described how they would handle each dilemma. Two of the scenarios were modified situations from Morris and Jacob’s (2000) past study. Another scenario
was inspired by the situations presented in Newman and Brown’s (1996) book. The final scenario was created based on the researcher’s previous evaluation experience.

Two forms to the SJT-E were created. The first (Appendix B) contained situations for judgement where no legitimate authority figure was present or making an unethical request. The second (Appendix C) form of the SJT-E used the same situations for judgement but were modified so a legitimate authority figure was present and making an unethical request. Responses to each of the situations were analyzed to determine the effect of authority figures on ethical decision making. A score which expressed the total sum of unethical decisions made was created for ease of examining total unethical options chosen on the SJT-E.

The following is an example of an SJT-E Authoritative Scenario: You are working for a philanthropic foundation, evaluating a program that they are a major sponsor for and fund. You, as the evaluator, have recently shared the draft of a final report with the director of the program being evaluated. After reviewing the draft, the program director asks the evaluator to tone down one section of the report that describes some operational problems within the program. The director believes that the findings in this section, although accurate, are presented in a way that could cause readers to overlook the overall success of the program’s implementation. You review the report one more time and the findings in the draft regarding operational problems appear to have been reported in a fair and balanced fashion. (Modified from Morris & Jacobs, 2000).

The following is an example of an SJT-E Non-Authoritative Scenario: You are working for a philanthropic foundation, evaluating a program that they are a major sponsor for and fund. You, as the evaluator, have recently shared the draft of a final report with your evaluation colleagues. After reviewing the draft, there has been talk among your evaluation colleagues of
toning down one section of the report that describes some operational problems within the program. The findings in this section, although accurate, are presented in a way that could cause readers to overlook the overall success of the program’s implementation. You review the report one more time and the findings in the draft regarding operational problems appear to have been reported in a fair and balanced fashion. (Modified from Morris & Jacobs, 2000).

In addition to the SJT, evaluators were asked to “please provide an explanation for your response to” each situation, in order for them to justify their response. These comments were analyzed to identify any mention of ethical codes, guidelines, principles, or major themes in rationale. Participants were also asked to rate “how ethically problematic Scenario [number]” was. Additionally, in order to check to see if an authoritative pressure was perceived, participants were asked to rate “how much of an authoritative presence” each scenario had. Finally, participants’ sex (male or female), employment status and tenure, and highest level of education were measured and assessed for possible interaction effects or moderation.

Data Collection Procedure

Two thousand AEA members were contacted to participate in the online situational judgement test and self-selected themselves into the study. Each participant was randomly presented with one of the two forms of SJT-E. The time it took to complete the SJT-E ranged from 10 to 40 minutes, depending on the thoroughness of responses to open-ended questions. Prior to access to the online SJT-E, participants were presented with an implied consent form, detailing anonymity, right to withdraw, that there are no benefits or consequences for participating, an overview of the purpose and the author, and the statement of approval from the IRB (see Appendix A). By clicking “next,” participants electronically agreed to the implied consent form and were taken to either the Non-Authority SJT-E (Appendix B) or the Authority
SJT-E (Appendix C). After submitting their responses online, participants received a final page that provides gratitude from the author, debriefing information, and contact information of the author in case additional questions about the study arise (Appendix D).

Data Analysis

Quantitative data was analyzed using IBM SPSS Statistics 23 software package. Qualitative comments from the rationale portion of the SJT-E forms were themed using QSR NVivo 11 software. A combination of Chi-Square, t-tests, and Pearson’s r analyses were used to test the study’s hypotheses.

**H1: Situations with authoritative pressure to comply with unethical requests (SJT-E: Authoritative) will illicit more unethical responses than situations without pressure to comply with unethical requests (SJT-E: Non-Authoritative).** Scores based on responses to the SJT-Es were compared using a Chi-Squared cross-tabulation matrix to test of scenarios in each form resulted in more or less ethically problematic responses.

**H2: Years as an evaluator practitioner will be related to amount of unethical responses.** A correlation between years in the evaluation practice and total scores on the SJT-E were assessed using a Pearson’s r analysis. A t-test was used to compare SJT-E forms on years as an evaluator practitioner and total amount of unethical responses made on the SJT-E.

**H3: Greater familiarity with ethical codes, standards, and principles will influence how ethical participant’s responses to scenarios will be.** An independent t-test on self-reported level of familiarity with ethical codes, standards, and principles between scores on each scenario was analyzed to assess if participants differed significantly.

Additional analyses. Differences between SJT-E forms and responses to rating how ethically problematic scenarios were and authoritative presence were assessed using an
independent t-test. Content analysis of participants’ rationale for each scenario choice by qualitative analysis coders were examined for major themes in explanations.

**Summary**

A situational judgement test regarding ethical decision making was used to examine the ethical decision making of evaluators in ethical dilemmas. Two forms of the test were created, with one having legitimate authoritative presence and the other having vague or illegitimate authoritative presence. An email invitation to the online SJT-E was sent to 2,000 current AEA members. A total of 243 AEA members provided usable, complete responses (12.15% response rate). Scores on the test and responses to qualitative and demographic items were analyzed to provide evidence of the current decision making and ethical code, standards, and principles use.
Chapter IV: Results

The purpose of this study was to assess the influence authority figures have on decision making in ethically problematic situations. Specifically, this study determines Milgram’s (1963) obedience to authority theory’s influence on professional evaluators using a situational judgement test (SJT-E). This section describes the data cleaning procedure, the manipulation check results, the results of the hypotheses testing by hypothesis, and the thematic analysis of the qualitative comments.

Data Cleaning

In order to ensure quality of responses to the SJT-E, blind coders were instructed to categorize the responses participants provided as rationale to scenario choices as being for the unethical response, against the unethical response, or ambiguous rationale to the scenarios. After un-hiding the actual response choice to scenarios, categorization was compared to response. Participants who provided rationale opposite to what they responded to the scenario with were excluded from the following analysis.

Participants who provided rationale that opposed their scenario choice or criticized the scenario in general were deemed problematic to the quality of data because the researcher could not be certain the participant responded with their intended choice. For example, rationale that does not align with response choice made could indicate the participant clicked the wrong choice by accident or did not fully understand the scenario or justification question. Additionally, participants who criticized the scenario only were deemed problematic because their true response to the scenario would not be determined. For example, one participant who criticized the scenario commented that they choose one response at random so they could move onto the next scenario without being stopped by the forced response warning in Qualtrics. For this reason,
it was unclear whether or not participants who criticized the scenario choose a response at random.

Scenario two had ten participants whose responses were removed because their rationale for the choice did not match their actual response chosen. Scenario Three had two participants whose response were removed because their rationale did not match the actual response chosen. Scenario Four had six participants whose response were removed because rationale did not match actual response chosen. A total of 18 participant’s responses were excluded from the following analysis after a blind coding. It is important to note that participants were not removed fully from all analyses for having hypocritical rationale or criticizing comments to one scenario. Instead, only their quantitative response to the scenario in question was recoded as missing. For example, only a participant’s problematic response to the scenario in question was removed as they responded typically to other scenarios and SJT-E items.

Manipulation Check: Authoritative Presence and Ethically Problematic Ratings

Participants were asked to rate the authoritative presence they felt regarding each scenario on a scale from one – no authoritative presence to four – a lot of authoritative presence. Likewise, participants rated the how ethically problematic each scenario was on a scale from one – not ethically problematic to four – very ethically problematic. Results of a $t$-test revealed significant differences between the no-authority ($M = 2.45, SD = 1.15$) and authority ($M = 2.89, SD = .9$) situational judgement test forms on Scenario Three authority presence ratings ($t(183.45) = -3.05, p = .003$, Cohen’s $d = .45, 95\%$ CI: -.72, -.15). No other significant differences were found between the two situational judgement test forms and ratings of authority presence. (see Table 4).
Results of a $t$-test suggested significant differences between situational judgement test forms and ratings of how ethically problematic scenarios were for Scenario One, Three, and Four (see Table 5). Participants who received the situational judgement test- authority form ($M = 2.60, SD = 1.06$) rated Scenario One as significantly more ethically problematic compared to participants who received the non-authority form ($M = 1.79, SD = .78$) ($t(234.05) = -7.11, p < .001$, Cohen’s $d = .93$, 95% CI: -1.02, - .58). Participants who received the authority form ($M = 3.06, SD = 1.05$) rated Scenario Three as more ethically problematic compared to participants who received the non-authority form ($M = 2.52, SD = 1.17$) ($t(222.46) = -3.74, p < .001$, Cohen’s $d = .50$, 95% CI: -.83, -.26). Likewise, participants who received the authority form ($M = 2.60, SD = .98$) rated Scenario Four as more ethically problematic compared to participants who received the non-authority form ($M = 1.89, SD = .97$) ($t(234) = -5.59, p < .001$, Cohen’s $d = .73$, 95% CI: -.96, -.46). Additionally, a Pearson’s $r$ correlation found authoritative presence and ethically problematic ratings for all scenarios were related (see Table 6).

Table 4

Authority Presence Ratings by Scenario

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Non-Authority Form</th>
<th>Authority Form</th>
<th>$t$</th>
<th>$df$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$n$</td>
<td>$M$</td>
</tr>
<tr>
<td>Scenario 1</td>
<td>2.20</td>
<td>.91</td>
<td>102</td>
<td>2.68</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>2.42</td>
<td>1.03</td>
<td>99</td>
<td>2.76</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>2.45</td>
<td>1.15</td>
<td>99</td>
<td>2.89</td>
</tr>
<tr>
<td>Scenario 4</td>
<td>2.05</td>
<td>.96</td>
<td>100</td>
<td>2.68</td>
</tr>
</tbody>
</table>
Table 5

*Ethically Problematic Ratings by Scenario*

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Non-Authority Form</th>
<th>Authority Form</th>
<th>t</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>n</td>
<td>M</td>
</tr>
<tr>
<td>Scenario 1</td>
<td>1.79</td>
<td>0.78</td>
<td>111</td>
<td>2.60</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>3.02</td>
<td>1.06</td>
<td>111</td>
<td>3.12</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>2.52</td>
<td>1.17</td>
<td>111</td>
<td>3.06</td>
</tr>
<tr>
<td>Scenario 4</td>
<td>1.89</td>
<td>.97</td>
<td>111</td>
<td>2.60</td>
</tr>
</tbody>
</table>

Table 6

*Correlations between Authoritative Presence and Ethically Problematic Ratings by Scenario*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Authoritative presence S1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Authoritative Presence S2</td>
<td>.30**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Authoritative Presence S3</td>
<td>.30**</td>
<td>.46**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Authoritative Presence S4</td>
<td>.39**</td>
<td>.34**</td>
<td>.41**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Ethically Problematic S1</td>
<td>.36**</td>
<td>.19**</td>
<td>.19**</td>
<td>.20**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Ethically Problematic S2</td>
<td>.17*</td>
<td>.45**</td>
<td>.39**</td>
<td>.16*</td>
<td>.25**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Ethically Problematic S3</td>
<td>.22**</td>
<td>.32**</td>
<td>.44**</td>
<td>.27**</td>
<td>.35**</td>
<td>.42**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8. Ethically Problematic S4</td>
<td>.16*</td>
<td>.28**</td>
<td>.27**</td>
<td>.51**</td>
<td>.37**</td>
<td>.31**</td>
<td>.45**</td>
<td>1</td>
</tr>
</tbody>
</table>

*Note. *p < .05, **p < .01*
Results by Hypothesis

The following section breaks down the analyses conducted for this study by hypothesis. Hypothesis testing included Chi-Square, Fisher’s Exact Test, t-tests, and Pearson’s r analyses. Following the hypothesis testing results, qualitative rationale for choice selection from Scenario Three is examined and discussed.

H1: Situations with authoritative pressure to comply with unethical requests (SJT-E: Authoritative) will illicit more unethical responses than situations without pressure to comply with unethical requests (SJT-E: Non-Authoritative). Participants who received the SJT-E Non-Authority form and participants who received the Authority form did not differ on their response to Scenario One ($X^2 (1) = 1.98, p = .159$). Of the participants who received the Non-Authority form, 57 chose the ethical option while 54 chose the unethical option. Seventy-five Participants who received the Authority form chose the ethical option to Scenario One while 49 chose the unethical option. Likewise, the SJT-E forms did not differ on participant responses to Scenario Two ($X^2 (1) = .60, p = .437$). Eighty-three participants who took the Non-Authority form chose the ethical response option while 29 chose the unethical response option for Scenario Two. Likewise, 98 participants who took the Authority form chose the ethical option in Scenario Two while 56 chose the unethical option.

Chi-Square analysis revealed that participants who received the Non-Authority SJT-E and the Authority SJT-E significantly differed on response to Scenario Three ($X^2 (1) = 6.97, p = .008$). Of the participants who received the Non-Authority form, 107 responded with the ethical option while six chose the unethical option. Of the participants who received the Authority form, 105 chose the ethical option while 20 chose the unethical option to Scenario Three. Results of an Odds Ratio suggest participants who receive the Authority form of Scenario Three were 3.40
times more likely to choose the unethical option than those who received the Non-Authority form of Scenario Three (95% CI: 1.31, 8.79) (see Table 7).

A Fisher’s Exact Test was used to examine the responses to Scenario Four between SJT-E forms. This analysis was chosen because less than five participants responded with the unethical option in Non-Authority form. The results of the Fisher’s Exact Test revealed no significant differences between the Non-Authority and Authority forms, (Fisher’s Exact = 1.07, p = .300). Of the participants who received the Non-Authority SJT-E, 111 responded with the ethical option and two responded with the unethical option. Of those that received the Authority SJT-E, 118 chose the ethical option while seven chose the unethical option in Scenario Four.

Table 7

*Scenario Three Responses by SJT-E Form*

<table>
<thead>
<tr>
<th>Scenario 3 Response Options</th>
<th>Non-Authority</th>
<th>Share the names</th>
<th>Total</th>
<th>$X^2$</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>SJT-E Form</td>
<td>Non-Authority</td>
<td>107</td>
<td>6</td>
<td>113</td>
<td>6.97**</td>
</tr>
<tr>
<td></td>
<td>Authority</td>
<td>105</td>
<td>20</td>
<td>125</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* **$p < .01$.**

**H2: Years as an evaluator practitioner will be related to amount of unethical responses.** A composite score was created to represent the total amount of unethical responses made. Responses to each scenario were scored as zero – correct/ethical response choice or one – incorrect/unethical response choice. The Score on SJT-E composite is the total number of unethical response choices made by the participant (out of four possible). Results of a Pearson’s r correlation revealed no significant relationship between the amount of years actively practicing evaluation (Years as an Evaluator, M = 12.58, SD = 9.64) and amount of ethical responses (Score on SJT-E, M = .82, SD = .78) (r = .09, p = .177). Form cases were split to assess if a
relationship existed within each form of the SJT-E. When looking at non-authority SJT-E responses only, no relationship was found between years as an evaluator and the total amount of unethical responses on the SJT-E Non-Authority ($r = .07$, $p = .513$). Likewise, when using only Authority SJT-E responses, no relationship was found between years as an evaluator and amount of unethical responses on the SJT-E Authority ($r = .12$, $p = .206$). Results of a t-test revealed no significant differences between Non-Authority SJT-E ($M = 13.01$, $SD = 10.72$) and Authority SJT-E forms ($M = 12.20$, $SD = 8.62$) and years as an evaluator ($t (201.10) = .62$, $p = .537$). Likewise, Non-Authority ($M = .83$, $SD = .70$) and Authority forms ($M = .81$, $SD = .85$) of the SJT-E did not differ significantly on the amount of unethical responses made on the SJT-E ($t (225.45) = .18$, $p = .854$).

**H3: Greater familiarity with ethical codes, standards, and principles will influence how ethical participants’ responses to scenarios will be.** Participants were asked to rate how familiar they are with the Institutional Review Board (IRB) approval process, the AEA’s Guiding Principles (2004), The Essential Competencies for Program Evaluators (Stevahn, King, Ghere, & Minnema, 2005), and the Program Evaluation Standards (Yarbrough, Shulha, Hopson, & Caruthers, 2011) on a 4-point Likert scale ranging from Not at all Familiar to Very Familiar. Participants were most familiar with the IRB process ($M = 3.50$, $SD = .78$) and least familiar with the Essential Competencies for Program Evaluators ($M = 2.01$, $SD = 1.01$). Participants were slightly familiar with the Program Evaluation Standards ($M = 2.26$, $SD = 1.12$) and moderately familiar with the Guiding Principles ($M = 3.06$, $SD = 3.06$).

Results of a $t$-test comparing response chosen on Scenario One on familiarity of the IRB approval process, the AEA Guiding Principles, the Essential Competencies for Program Evaluators, and the Program Evaluator Standards revealed no significant differences (see Table
8). Results of a *t*-test comparing response chosen on Scenario Two on familiarity of the IRB approval process, the AEA Guiding Principles, the Essential Competencies for Program Evaluators, and the Program Evaluator Standards also revealed no significant differences (see Table 9). Likewise, results of a *t*-test comparing response chosen on Scenario Four on familiarity of the IRB approval process, the AEA Guiding Principles, the Essential Competencies for Program Evaluators, and the Program Evaluator Standards also revealed no significant differences (see Table 11).

Results of a *t*-test suggest response chose to Scenario Three did not significantly differ on familiarity with the AEA Guiding Principles, the Essential Competences for Program Evaluators, or the Program Evaluation Standards (see Table 10). However, participants who did not share the teachers names with program staff (*M* = 3.56, *SD* = .74) reported being more familiar with the IRB approval process compared to those who chose to share the teachers’ names (*M* = 3.00, *SD* = .98) (*t*(231) = -3.39, *p* = .001, Cohen’s *d* = .45, 95% CI: -.89, -.24).
Table 8

*Scenario One Responses and Familiarity with Ethical Codes, Standards, and Guidelines*

<table>
<thead>
<tr>
<th>Scenario 1 Response Options</th>
<th>You revise the section</th>
<th>You keep the section as is</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>t</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>The process of gaining approval from the IRB</td>
<td></td>
<td></td>
<td>3.55</td>
<td>0.72</td>
<td>103</td>
<td>3.46</td>
<td>0.83</td>
<td>129</td>
<td>.93</td>
<td>230</td>
</tr>
<tr>
<td>The AEA’s Guiding Principles</td>
<td></td>
<td></td>
<td>3.15</td>
<td>0.89</td>
<td>103</td>
<td>2.99</td>
<td>0.99</td>
<td>129</td>
<td>1.23</td>
<td>230</td>
</tr>
<tr>
<td>The Essential Competencies for Program Evaluators</td>
<td></td>
<td></td>
<td>2.05</td>
<td>1.11</td>
<td>102</td>
<td>2.01</td>
<td>0.92</td>
<td>127</td>
<td>.30</td>
<td>4</td>
</tr>
<tr>
<td>The Program Evaluation Standards</td>
<td></td>
<td></td>
<td>2.34</td>
<td>1.14</td>
<td>103</td>
<td>2.22</td>
<td>1.10</td>
<td>128</td>
<td>.82</td>
<td>229</td>
</tr>
</tbody>
</table>

Table 9

*Scenario Two Responses and Familiarity with Ethical Codes, Standards, and Guidelines*

<table>
<thead>
<tr>
<th>Scenario 2 Response Options</th>
<th>Use of passive consent</th>
<th>Use of active consent</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>t</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>The process of gaining approval from the IRB</td>
<td></td>
<td></td>
<td>3.62</td>
<td>.60</td>
<td>50</td>
<td>3.47</td>
<td>.83</td>
<td>174</td>
<td>1.46</td>
<td>0</td>
</tr>
<tr>
<td>The AEA’s Guiding Principles</td>
<td></td>
<td></td>
<td>3.04</td>
<td>.86</td>
<td>50</td>
<td>3.04</td>
<td>.97</td>
<td>174</td>
<td>-0.00</td>
<td>222</td>
</tr>
<tr>
<td>The Essential Competencies for Program Evaluators</td>
<td></td>
<td></td>
<td>2.04</td>
<td>1.00</td>
<td>49</td>
<td>2.02</td>
<td>1.01</td>
<td>172</td>
<td>.107</td>
<td>219</td>
</tr>
<tr>
<td>The Program Evaluation Standards</td>
<td></td>
<td></td>
<td>2.08</td>
<td>.99</td>
<td>50</td>
<td>2.30</td>
<td>1.15</td>
<td>173</td>
<td>-1.34</td>
<td>90.84</td>
</tr>
</tbody>
</table>
Table 10

*Scenario Three Responses and Familiarity with Ethical Codes, Standards, and Guidelines*

<table>
<thead>
<tr>
<th>Scenario 3 Response Options</th>
<th>Share the names</th>
<th>Do not share the names</th>
<th>t</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>The process of gaining</td>
<td>M</td>
<td>SD</td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>approval from the IRB</td>
<td>3.00</td>
<td>.98</td>
<td>24</td>
<td>3.56</td>
</tr>
<tr>
<td>The AEA’s Guiding Principles</td>
<td>2.71</td>
<td>1.12</td>
<td>24</td>
<td>3.10</td>
</tr>
<tr>
<td>The Essential Competencies for Program Evaluators</td>
<td>2.13</td>
<td>1.04</td>
<td>24</td>
<td>1.99</td>
</tr>
<tr>
<td>The Program Evaluation</td>
<td>1.88</td>
<td>.85</td>
<td>24</td>
<td>2.30</td>
</tr>
</tbody>
</table>

Table 11

*Scenario Four Responses and Familiarity with Ethical Codes, Standards, and Guidelines*

<table>
<thead>
<tr>
<th>Scenario 4 Response Options</th>
<th>Include the art guild</th>
<th>Do not include the art guild</th>
<th>t</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>The process of gaining</td>
<td>M</td>
<td>SD</td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>approval from the IRB</td>
<td>3.50</td>
<td>.79</td>
<td>225</td>
<td>3.33</td>
</tr>
<tr>
<td>The AEA’s Guiding Principles</td>
<td>3.06</td>
<td>.94</td>
<td>225</td>
<td>3.67</td>
</tr>
<tr>
<td>The Essential Competencies for Program Evaluators</td>
<td>2.02</td>
<td>1.01</td>
<td>223</td>
<td>3.00</td>
</tr>
<tr>
<td>The Program Evaluation</td>
<td>2.28</td>
<td>1.13</td>
<td>225</td>
<td>2.00</td>
</tr>
</tbody>
</table>
Rationale for Scenario Three Thematic Analysis

Participants provided qualitative responses to the question “please provide an explanation for your response” to each of the SJT-E scenarios. Content analysis was completed by three qualitative analysis coders to systematically reduce participant responses into categories then themes to inductively analyze the overall story of the data. Responses were not altered for grammar, spelling, or content. Qualitative analysis coders independently themed the rationales for all scenarios. After a discrepancy meeting, coders settled on set theme names and definitions then independently completed a second coding with the agreed upon theme structure. After a final discrepancy meeting between coders, a minimum of 95% coding comparison agreement was met. The following section describes the themes that emerged out of Scenario Three’s content analysis only as Scenario Three was the only scenario in which significant differences between SJT-E forms were found. The theme tables for Scenarios One, Two, and Four can be found in Appendices E-J along with a description of the thematic analysis.

Of the participants who received the Non-Authority SJT and provided rationale for Scenario Three, the majority chose to not share the teachers’ names. Many participants commented on providing “Alternative Actions” that could be taken during data collection to improve the response rate and avoid having to break confidentiality. Participants also rationalized not sharing names because doing so would break participant “Confidentiality”. Of the participants who received the Authority SJT-E and provided rationale for Scenario Three, the majority also chose not to share teacher’s names. Participants also rationalized that sharing the names would break “Confidentiality” and go against “Ethics”. Of the participants who took the Authority SJT-E and provided rationale for providing the names, many comments suggested the “Client” was a factor in their decision making and their situation “Depends” on other factors.
Table 12

*Scenario Three Rationale Themes: Non-Authority SJT-E*

<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does not share teacher’s name</td>
<td>Subthemes include: Confidentiality, Alternative Action, Ethics, Disciplinary, Informative, and Other</td>
<td></td>
<td>- I think instead I would intervene in moment to ask folks to complete post items. My goal is to increase participation.</td>
</tr>
<tr>
<td></td>
<td>Participants suggested an alternative response to the scenario</td>
<td>28</td>
<td>- I would have tried to get those teachers back on task at the time of data collection.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- I would ask the teachers to complete the evaluation items while I still had the opportunity to get the data. That is what is important, not &quot;catching&quot; the teachers.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- I assume the participants were offered confidentiality. I would note my observation but not identify individuals.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- This would be a breach of confidentiality.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Reporting the teachers will violate their confidentiality and be against voluntary participation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- That action would be unethical in that it crosses the line between program evaluation and supervision.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Clearly this is an ethical issue. By telling the supervisor who these teachers are you open them up to retribution or potential risk of some sort of harm.</td>
</tr>
<tr>
<td>Confidentiality</td>
<td>Sharing the names would break participant confidentiality</td>
<td>27</td>
<td>- It would be unethical to share the names.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Clearly this is an ethical issue. By telling the supervisor who these teachers are you open them up to retribution or potential risk of some sort of harm.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- That is what is important, not &quot;catching&quot; the teachers.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- It benefits nobody to turn this evaluation into an accountability and possibly disciplinary situation.</td>
</tr>
<tr>
<td>Alternative Action</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethics</td>
<td>Respondents comment that it would be unethical to provide teachers names</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Disciplinary</td>
<td>Respondents feel it is not their place to &quot;narc&quot; or provide the names of teacher because non participants may be disciplined</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Theme</td>
<td>Definition</td>
<td>Count</td>
<td>Examples</td>
</tr>
<tr>
<td>----------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Informative</td>
<td>The level of participation or lack of is informative on its own</td>
<td>10</td>
<td>• I am assuming there is no contractual obligation to report individual teacher behavior. The level of participation in the survey is in fact informative.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Further, non-completion is an evaluation problem, not a program problem. It appears that the process for data collection did not induce cooperation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• lack of participation in an actively is data. What is needed is to figure out the meaning of the response or lack of. For example, for those who may view the training as an imposition they may not be so inclined to respond to post training survey. This is more important to know than to try to force people to respond.</td>
</tr>
<tr>
<td>Share the teacher’s names</td>
<td>Participants’ who responded in favor of sharing the non-participating teachers’ names</td>
<td>5</td>
<td>• Presuming this was a required teacher activity, not a voluntary one, it is acceptable to report on the teachers who were not fulfilling their work responsibilities.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• I think if people did not do it they should be held accountable and you need to address that your post-program data is not complete.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Unless the pre-program data collection process identified that this would be a totally anonymous process, as an evaluator I would feel obligated to report the names. If the school district is going to get the expected outcomes, they need the implementation process to be supported.</td>
</tr>
<tr>
<td>Other</td>
<td>Other comments that did not fall into an above subtheme</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Theme</td>
<td>Definition</td>
<td>Count</td>
<td>Examples</td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>-------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Do Not Share the Names</td>
<td>Subthemes include: Confidentiality, Ethics, Alternative Action, Disciplinary, Responsibility, and Other</td>
<td></td>
<td>- You do not indicate whether the teachers were promised confidentiality in their responses. If so, the answer is a definite DO NOT.</td>
</tr>
<tr>
<td>Confidentiality</td>
<td>Sharing the names would break participant confidentiality</td>
<td>55</td>
<td>- Confidentiality is important and if it was understood from the beginning that all surveys would be anonymous then that is how it should stay.</td>
</tr>
<tr>
<td>Ethics</td>
<td>Respondents comment that it would be unethical to provide teachers names</td>
<td>22</td>
<td>- The responses are supposed to be kept confidential, which means that participation on a particular evaluation activity should also be kept confidential.</td>
</tr>
<tr>
<td>Alternative Action</td>
<td>Participants suggested an alternative response to the scenario</td>
<td>15</td>
<td>- Ethical considerations require that the information applied to the scenario is untraceable and unbiased. Unfortunately, the program director can not get this information from the project team. The best you can do is give him a summary of the overall response rate.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- The survey could be re-administered after a &quot;pep talk&quot; to boast participation and commitment to the process.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- So, in that instance, I don't share the names or non-response data, but I do aggressively track the non-responders myself and try to get the data. And, I might ask the program director to send out an email to everyone</td>
</tr>
<tr>
<td>Theme</td>
<td>Definition</td>
<td>Count</td>
<td>Examples</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------------------------------------------------</td>
<td>-------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Disciplinary  | Respondents feel it is not their place to "narc" or provide the names of teacher because non participants may be disciplined | 12    | • However, I may make an announcement during the test and survey administration that we view their voices and feedback as very important components of the evaluation to encourage a better response rate.  
• Identifying information could put the teachers at risk, including those that answered. Better to look for another way to manage the issue -qualify the results, find a way to read-minister the questionnaire, etc.  
• It's unclear what the program director's motive is for requesting this. If it's to follow up with the non-respondents, I can do that just as well or better than s/he can. If it's it slap people on the wrist for not responding, that's unacceptable. Either way, I don't share the list.  
• Sharing the names of the teachers may result in reprisals such as dismissal. The names of survey participants should be kept confidential to prevent such adverse consequences. |
| Responsibility| Respondents’ comments regarding their role as an evaluator and the responsibilities and competencies they have as evaluators | 7     | • I'm going to assume that the evaluation contract doesn't include providing the identifiable data to the program director, so I will not provide this information.  
• The role of the evaluator is to evaluate outcomes, not teacher behavior. If the program director believes the outcomes are related to teacher behavior, then he/she can request a process evaluation.  
• I have to protect my integrity as an evaluator. |
<p>| Other         | Other comments that did not fall into an above subtheme                     | 15    |                                                                                                                                                                                                         |
| Share the Names | Subthemes include: Client, Depends, Participation is Mandatory,            |       |                                                                                                                                                                                                         |</p>
<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Follow-up, and Other | Pleasing the client and their requests is important |       | - As long as there was nothing said about anonymity, since this was a form of professional development and the program director is entitled to know who did or did not complete the assignment  
- As an external, independent evaluator, it is important to put the client first. In this case, the client is the program director.  
- Given that the program director is my client in this case, I'm behoven to provide the program director with the info that is requested as a part of the evaluation. |
| Client   | 6                                                                           |       |  
|          | Reporting the names of teachers who did not participate is a way to follow up | 4     | - Participation is required and this does not threaten confidentiality of the questions. No different than a follow-up request for a survey.  
- Completing the post-program evaluation is a reasonable request of participants and it is reasonable for the program director to follow-up with participants that did not submit the evaluation.  
- Whether or not to share the teachers’ names depends on other factors in the situation |
| Depends  | 6                                                                           |       | - As long as there was nothing said about anonymity, since this was a form of professional development and the program director is entitled to know who did or did not complete the assignment  
- I would share if and only if the teachers were aware of the lack of anonymity, and how the data would be used. I would also let the director know that if participants are aware that their data can be linked/identified, that might compromise valuable feedback with more muted responses. And I might on the spot have the director make an announcement to encourage a better response atmosphere.  
- If there's some reason that the participant ID numbers and teacher names shouldn't be linked, then I would not do this. But for a diversity training, there doesn't not seem to be a need for confidentiality (regarding their names, not their responses). |


<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Other</td>
<td></td>
<td>Other comments that did not fall into an above subtheme</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>• No different than a follow-up request for a survey.</td>
</tr>
</tbody>
</table>

**Summary**

Results of hypothesis testing revealed significant differences between SJT-E forms for Scenario Three. Additionally, familiarity with the IRB approval process differed significantly between Scenario Three response chosen. Thematic analysis of Scenario Three suggests participants who received the Authority SJT-E form and chose to provide non-participating teacher’s names to the client used the client’s presence as rationale for breaking confidentiality. Results of the manipulation check questions revealed authority presence and ethically problematic ratings of scenarios were significantly related. Finally, there were no evidence to support that years spent practicing evaluation as being correlated to the amount of unethical responses made on the SJT-E forms.
Chapter V: Discussion, Conclusion and Recommendation

The purpose of this study was to assess the influence obedience to authority has on the decisions evaluation practitioners make in ethically problematic situations. Additionally, the relationship between years spent as an evaluator and the total amount of unethical responses made on the SJT-E were examined. Finally, familiarity with ethical codes, standards, and guidelines were assessed for effect on scenario response made. The following section describes the implications of the hypotheses tested, limitations of the study, future research, and recommendations for practice.

Discussion

Results of this study provided partial support that authoritative presence influences decisions made in ethical dilemmas. Specifically, the results provided partial support that the Authority SJT-E form resulted in more unethical choices than the Non-Authority SJT-E form. While frequency of unethical responses did not differ significantly between SJT-E forms for scenarios One, Two, and Four, Scenario Three revealed that participants who received the authoritative SJT-E form were over three times more likely to share the names of teachers who did not participate in the post evaluation survey than those who received the non-authoritative form. Rationale for scenario three differed between forms. Specifically, comparing rationale for sharing teachers’ names suggests having a client present in the scenario three resulted in different major themes present. For example, one participant who took the authority SJT-E rationalized sharing the names as, “as an external, independent evaluator, it is important to put the client first. In this case, the client is the program director.” The differences between SJT-E forms and rationale for scenario three is partial evidence that authority presence influences the decisions made in ethically problematic situations.
Interestingly, participants who received the Authority SJT-E form rated scenarios One, Three, and Four as more ethically problematic compared to those who received the Non-Authority SJT-E form. Considering there were no significant difference between SJT-E forms and decisions chosen for Scenarios One and Four, it is notable that participants recognized those scenarios are being more ethically problematic when a client or stakeholder were present. Simply, the participants recognized the ethical dilemma and responded ethically regardless of client or stakeholder (authority figure) pressure.

There was no support for hypothesis two from the study, that the sum of unethical choices made would be related to the amount of time spent actively practicing evaluation. Interestingly, SJT-E forms did not differ significantly on the number of unethical responses made, suggesting that authoritative presence did not have an affect overall on decisions made. It is likely that the ethical, or “correct”, choice was apparent in each scenario, regardless of client presence. It may also be likely that American Evaluation Association (AEA) members are familiar enough with ethical codes and standards to make the most ethical choice in each situation. Likewise, asking how a participant should respond in the scenario may make it easier to choose the ethical response compared to real-life situations and pressures.

The study’s results revealed partial support for hypothesis three. Specifically, there was no significant difference between scenario responses to the SJT-E on familiarity with the Institutional Review Board (IRB) approval process, the AEA’s Guiding Principles (2004), The Essential Competencies for Program Evaluators (Stevahn, King, Ghere, & Minnema, 2005), and the Program Evaluation Standards (Yarbrough, Shulha, Hopson, & Caruthers, 2011) for scenarios one, two, and four. However, participants who chose not to share teachers’ names in scenario three reported being more familiar with the IRB process than those who shared the
teachers’ names. Rationale from participants who did not share the names, regardless of SJT-E form, suggested that confidentiality and ethics were major themes. These results suggest understanding and familiarity with IRB confidentiality and anonymity may play a role in making ethical decisions relating to identifying participants.

The IRB as an authority in decision making. What is interesting, regarding Scenario Three, is that the IRB and confidentiality was mentioned as rationale for and against sharing non-participating teacher’s names. The rationale comments suggest that the IRB is being used as an authority for the evaluation, when this line of thinking could potentially be detrimental. The IRB process is not unified across institutions and the regulations set forth by the IRB can differ depending on where approval is being processed (Stark, 2014). Stark (2014) wrote about the issue with IRB members using local precedent to make decisions regarding research, leading to differences in protocols across IRBs. Specifically, each IRB uses their knowledge of past research to make decisions about current proposals, thus differences develop between IRBs form.

Obedience to authority. Milgram’s (1974) obedience to authority experiments found that in extreme ethical dilemmas, participants will obey the experimenter, the authority in the situation. The same rate of obedience was not found in this study and could be explained by various differences. In particular, Milgram’s (1974)’s shock experiments placed participants in intense situations that involved life and death of a confederate, so the stakes and pressure on the participants differed greatly compared to the scenarios presented in the SJT-E. The power of the situation in each scenario reflects different pressures, not merely authoritative.

External and internal pressures. Motivations and values influence decisions as an internal source of pressure. External pressure, outside of authority presence, could also include pressure from peers, such as social norms. Overall, the scenarios in the SJT-E acted as a means
to compare client authority to non-client authority situations without being directly comparable to Milgram’s (1974) shock experiments. While obedience to authority may affect decisions made, it is unlikely to be the only pressure influencing the individual. Further research on ethical decision making and the power of the situation is necessary.

*The evaluator as an authority figure.* In addition, it is important to note that the authority figure present in an evaluation may be more ambiguous than it looks on the face of a situation. There is a difference between an evaluation practitioner’s self-authority and the authority of their client. Specifically, evaluators may see themselves as the authority in situations relating to evaluation and outcomes, rather than the client or stakeholder. This self-identification as the authority may explain the confusion surrounding the manipulation check question.

**Limitations and Future Research**

The limitations of the study bring to light areas of improvement for future studies on ethical decision making and research on evaluators. Due to the small sample size, the study’s results cannot be generalized outside of United States AEA members to the other 50 countries that home AEA Members. While a lack of representativeness of U.S. members is a limitation, it was not the intention of the study to examine obedience to authority in other countries. However, the low response rate (12%) was unexpected and suggests that the method of recruiting and inviting participants was not best practice for the topic of study. Participants were invited via email and told the SJT-E could take up to 40 minutes, depending on how thorough they were with responses. The long time estimate may have prevented many participants from even opening the SJT-E. Participants who are willing to take a situational judgement test may also differ from those who are unwilling to participate. It may also be a possibility that because the
study took place immediately after the 2015 AEA conference that participants were already being oversampled and burnt out.

Additionally, while a random selection of 2,000 AEA members was provided for the study, true random selection was not obtainable. Selection bias is an area of concern because participants self-selected into the study. However, random assignment to conditions was possible through mathematical number generation. Specifically, the random sample of 2,000 participants were assigned a random number using a random number generator, then sorted by ascending values. The top 1,000 participants were selected into the non-authority SJT-E condition while the remaining 1,000 were selected into the authority SJT-E condition.

Another limitation of the study regards the ratings of authoritative presence in each scenario. Participants were given the option to provide additional comments and the majority used this item to comment on not understanding what authoritative presence meant. While confusion and low ratings were expected for the non-authority SJT-E form, many participants, regardless of SJT-E form, found the question confusing. Comments suggested participants thought they should rate their own authoritative presence or the authority presence in general. Therefore, the results of the authoritative presence ratings should be interpreted with caution. Future research should rephrase this manipulation check question or provide additional explanation of authoritative presence so participants have a clear understanding of what is being asked.

Participants’ responses were also cleaned from the dataset if their response to the rationalization question after each scenario were hypocritical to their actual scenario response chosen or were overly critical of the scenario and its choices. This is a limitation of the study, as many participants felt strong enough about the scenarios that, instead of following question
directions, responded that the scenario lacked enough information to make a comfortable choice or that the lack of an alternative, middle-ground response option was unsettling. While, in the real world, situations have much more information to take into consideration while making decisions, this study was purposeful in making the scenarios ambiguous enough that the correct or ethical choice would be less obvious. Additionally, while responses to scenarios are not always either unethical or ethical, middle-ground alternatives are not always fully compliant with ethical codes, standards, and guidelines. Simple compromises can seem like a good alternative, but can also be considered unethical depending on the dilemma under different lenses.

The use of self-report data on an ethical decision making measure is another limitation of this study. Social desirability and exaggeration effects may be present which decreases the validity of the results (Wouters, Maesschalck, Peeters, & Roosen, 2014). Scenarios two and four resulted in the majority of participants choosing the ethical option, which could be attributed to the scenarios having an obvious ethical or unethical response set. While it is likely that participants who chose the ethical option understand and would make the same choice in a real situation, there is also the possibility participants recognized one option was more ethical than the other and chose their response based on a desire to be socially acceptable or “good” (Wouters, et al., 2014). Likewise, participants may exaggerate on how ethical they would be in real life (Randall & Fernandes, 1991). Stanley Milgram (1974) had originally expected low rates of obedience to authority based on initial surveying of psychologists and colleagues. Milgram (1978) suggests individuals may respond with what they know is the ethical choice but behave counter in an actual situation.

Additionally, a self-report test using only four scenarios is not a comprehensive measure of ethical decision making as an observational or qualitative study. Reliability of the SJT-E
would improve if more scenarios were added to the measure and if current scenarios were revised based on this study’s findings regarding ethically problematic ratings. A larger situational judgement test should be developed to obtain more reliable results and valid interpretations.

Lastly, future studies should take into consideration the definition of ethics and the best method of assessing ethical behaviors. Ethics is a cultural concept, and therefore the “correct” decision in any situation differs depending on the lens through which we look at the dilemma. Specifically, a country’s subjective well-being, economic development, and financial stability are all moderating variables when studying ethics across cultures (Wang & Murnighan, 2014). While life would be easier if the world agreed collectively on ethical standards, ethical judgements and perceptions differ based on the values of where they live and the overall well-being of their lives (Wang & Murnighan, 2014).

A potential future study could ask participants to closely journal their days over a course of an evaluation project and use the journal responses to identify ethically problematic situations and the decisions the participant made during. Alternative options for studying ethical decision making without the use of self-report measures could include observational designs in which evaluators are systematically observed and coded based on real situations. However, individuals may be reluctant to be observed or measured in ethical contexts, as anonymity is lost and fear of incriminating themselves or others is a potential risk (Randall & Fernandes, 1991). According to Randall and Fernandes (1991), using anonymous testing conditions or including the Balanced Inventory of Desirable Responding (BIDR) social desirability scale to assess bias during data collection would be an appropriate way to study ethics and obtain honest responses.
Implications for Practitioners

This study suggests practitioners should keep in mind the influence client and stakeholder relationships have on their perceptions of situations and the decisions they make. Many personal blogs, journal articles, webinars, and workshops have been devoted to talking about ethics in project settings. Practitioner would benefit from the development of a standard method or strategy for dealing with ethical dilemmas. For practitioners in the evaluation field, this study proposes the use of McDonald and Myrick’s (2008) method of identifying the advantages and disadvantages of different decisions in ethically problematic situations. The table in McDonald and Myrick’s (2008) article provides a strategy to systematically analyze the stakeholders and options in the situation (see Table 14). McDonald and Myrick (2008) also suggest that reflective practice is a means to learn from ethical challenges to make better decisions and prevent dilemmas from occurring in the first place.

<table>
<thead>
<tr>
<th>Stakeholders</th>
<th>Option 1</th>
<th>Option 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Advantages</td>
<td>Disadvantages</td>
</tr>
</tbody>
</table>

*Figure 1. McDonald & Myrick ethical decision making table. (2008)*

Additionally, Novo Nordisk and the University of Copenhagen developed an Ethics Dilemma and Ethics Decision-Making Tool that assists in making ethical decisions in grey zones. This tool has participants answer nine questions regarding the compliant, responsibility, and general acceptability of the decision in a checklist form (Copenhagen University, Novo Nordisk, & LRN, n.d.). While the tool was originally developed by pharmaceutical researchers, it is transdisciplinary in use (see Table 15). Evaluation Practitioners would also find the decision making tool checklist useful when dealing with ethical dilemmas in grey areas.
<table>
<thead>
<tr>
<th>Is it complaint:</th>
<th>Is it legal?</th>
<th>Will I break any rules or regulations?</th>
<th>Is it against company or organization policies?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is it responsible:</td>
<td>How does it affect other parties?</td>
<td>What if everybody was doing it?</td>
<td>Is it against my principles?</td>
</tr>
<tr>
<td>Is it generally acceptable:</td>
<td>Would it be okay to see it on the front page of a newspaper?</td>
<td>Would I be ashamed if my peers knew about it?</td>
<td>Could I tell my family about it?</td>
</tr>
</tbody>
</table>

Figure 2. Novo Nordisk & University of Copenhagen ethics decision-making tool.

**Recommendations for Education**

Education regarding obedience to authority, and in particular the relationship between consultants and clients, is important in preparing new and current evaluators for ethical dilemmas. Institutions and the AEA have a responsibility to both students and AEA members to provide adequate training on ethical decision making. Specifically, providing training on ethics and professional development on Institutional Review Board protocols is essential.

**Current student evaluators.** Slocum, Rohlfer, and Gonzalez-Canton (2014) used micro-insertion techniques to insert ethics training into regular coursework combined with external environmental practice. The same technique can be applied to graduate student evaluators or evaluators in training (Slocum, et al., 2014). Specifically, teachers of evaluation should introduce ethical issues frequently throughout their regular curriculum as part of a social consciousness activity. Additionally, the real-world applications of evaluation and interactions with clients are beneficial for students to discuss ethical dilemmas that occur when working with actual clients. These micro-insertions of ethical topics and real-world experience combined could provide better preparation of students to become ethical and experienced practitioners.
**Current evaluation practitioners.** Professional development for current evaluators would be beneficial to increasing familiarity and understanding of the multiple ethical standards, codes, and guidelines. For example the field should have discussions about making ethical decisions when faced with challenges and client/stakeholder pressure. In particular, training can be provided as a conference activity by specialists in workplace ethics or as a webinar introducing e-learning platforms such as Ethics dilemma and Ethics Decision-Making program by Copenhagen University, Novo Nordisk, and LRN (Copenhagen University, Novo Nordisk, & LRN, n.d.).

**Conclusions**

Overall, Milgram’s (1974) obedience to authority theory has some presence in evaluation, but it is unlikely that it is the only role that influences decision making in ethical dilemmas. The majority of participants, regardless of SJT-E form, chose the ethical response to all of the scenarios, suggesting that the clear ethical choice was recognized. However, in real life, ethical dilemmas are less ambiguous than those in the scenarios, authority can be diffused across individuals present, and there are more than two responses. No significant findings were identified for scenarios One, Two, and Four between SJT-E forms. However, Scenario Three is evidence that discussion as well as training regarding confidentiality in research and the IRB would be beneficial for the field. It is also clear that evaluation practitioners should be aware of the power of the situation. Specifically, evaluators should always keep in mind how the presence of clients and stakeholders can influence the decisions they make in potentially ethically problematic situations. This study contributes to the body of research on ethics in evaluation and obedience to authority in applied settings. Additionally, this study highlights areas in which the field of evaluation can improve. Future research should focus on improving the methodology
used to study ethical decision making and study the power of the situation on evaluators’ choices in ethically problematic situations. As a field, evaluators should start a larger discussion regarding how to remain mindful of both moral obligations as well as ethical obligations in a field where these decisions may not be as clear as they initially seem.
References


Appendix A: Implied Consent

The following questionnaire is a situational judgement test, using written scenarios as prompts. Each scenario will describe a potentially problematic situation that evaluators have faced. There will be a list of choices to choose from following each scenario. Please respond with the choice that best represent what you would suggest as a solution in the situation. The results of this situational judgement test will be presented in a thesis report for credits. Your responses to each question is confidential. It is believed that you cannot be identified by any of the following questions.

Your participation in the study is voluntary, and you may skip questions or close the website at any time without consequence. You will not receive any benefits for participating in the study, but your participation is highly appreciated by the thesis student.

This study has been reviewed and approved by University of Wisconsin-Stouts’ Institutional Review Board (IRB). The IRB has determined that this study meets the ethical standards required by federal law and University policies. This study was also reviewed and approved by the American Evaluation Association (AEA). Any questions or concerns related to the study should be directed towards the thesis student, Julia Lamping, or the assisting professor, Tiffany Smith. Any questions or concerns regarding your rights as a participant, contact the IRB Chair, Nels Paulson.

<table>
<thead>
<tr>
<th>IRB Chair</th>
<th>Thesis Student</th>
<th>Assisting Professor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nels Paulson</td>
<td>Julia Lamping</td>
<td>Dr. Tiffany Smith</td>
</tr>
<tr>
<td><a href="mailto:paulsonne@uwstout.edu">paulsonne@uwstout.edu</a></td>
<td><a href="mailto:lampingj0825@my.uwstout.edu">lampingj0825@my.uwstout.edu</a></td>
<td><a href="mailto:smithtiff@uwstout.edu">smithtiff@uwstout.edu</a></td>
</tr>
<tr>
<td></td>
<td>262.751.8211</td>
<td>(715) 232-2647</td>
</tr>
</tbody>
</table>

By clicking “NEXT” you are consenting to participation in this study.
Appendix B: SJT-E: Non-Authority

SJT-E: Illegitimate or No Authority

INSTRUCTIONS:
- This study is designed to take XX minutes
- For each SITUATION choose ONE from XX possible responses, which address the situation when done together. You must only select the ONE response that BEST describes how you would address the situation.
- Answer what you should do as an evaluator practitioner.
- You may sometimes feel you would like more information before answering, but please answer each question based only on the information provided.

1. THE REVISED REPORT SCENARIO
You are working for a philanthropic foundation, evaluating a program that they are a major sponsor for and fund. You, as the evaluator, have recently shared the draft of a final report with your evaluation colleagues. After reviewing the draft, you see that one section of the report that describes some operational problems within the program. The findings in this section, although accurate, are presented in a way that could cause readers to overlook the overall success of the program’s implementation. You review the report one more time and the findings in the draft regarding operational problems appear to have been reported in a fair and balanced fashion.

Response options:
- You revise the section in question.
- You keep the report as is and disseminate it to the client.

Please provide an explanation for your response to Scenario 1. <text box>

2. THE PASSIVE CONSENT SCENARIO
In evaluating a school-based youth program which educates students on safe sex at an early age, you have the choice of using either an active-consent or a passive-consent procedure to obtain parental permission. Active consent requires parents to sign and return a form if they wish to give permission for their child to participate in a study. In contrast, passive consent only requires them to sign and return a form if they do not want their child to participate. In general, it is much easier to achieve high participation rates with passive-consent approaches than with active-consent ones. Parents may not allow their children to participate in the survey; which, asks students about their sexual behaviors. Using passive consent will lead to a higher response rate because some parents who oppose the study will simply forget to return the passive-consent form, whereas others who would have opposed the study will fail to read the form in the first place.

Response Options:
- You decide to use the passive consent procedure.
- You decide to use the active consent procedure.

Please provide an explanation for your response to Scenario 2. <text box>
3. THE CONFIDENTIALITY AND PUNISHMENT SCENARIO
You are evaluating a diversity training program’s outcomes. The program is utilized throughout school districts as a form of teacher professional development in creating a diverse and safe classroom. The evaluation planning runs smoothly and you are even able to collect pre-program data before the training is implemented at the school district. A summer meeting is organized with all the teachers in the school district as a date to collect post-program data. You administer the post-program diversity test and accompanying survey. As you walk around the room, you notice not all the teachers are working on the evaluation item; but rather, they are using this time to socialize and disrupt those who are working. The participant ID numbers are linked to the teachers’ names in a separate excel workbook, which means you have the names of the teachers who do not complete the post-program evaluation items. You are able to identify and report the teachers who did not complete their post-program evaluation items to the program director.

Response Options:
- You share the names of teachers who did not participate in the post-program evaluation items.
- You do not share the name of the teachers who did not participate in the post-program evaluation items.

Please provide an explanation for your response to Scenario 3. <text box>

4. THE EVALUATION COMMITTEE EXCLUSION SCENARIO
You are conducting a summative evaluation of a community arts program, which is funded through government grants. Key stakeholders include local grant foundations, the community agency who is implementing the program, the local arts guild, the local community center where the program is run, and a few other community organizations that specialize in the arts. You plan to form an evaluation committee out of representatives of the stakeholders so that differing opinions on the evaluation focus and planning can be heard throughout the process. Through the grape-vine, you hear that the local arts guild is known in the community as being against formal assessments of any kind.

Response Options:
- Form the evaluation committee with representation from the local arts guild.
- Form the evaluation committee without representation from the local arts guild.

Please provide an explanation for your response to Scenario 4. <text box>

Rating Scenarios 1-4:

Scenario 1
You are working for a philanthropic foundation, evaluating a program that they are a major sponsor for and fund. You, as the evaluator, have recently shared the draft of a final report with your evaluation colleagues. After reviewing the draft, you see that one section of the report that describes some operational problems within the program may need toning down. The findings in this section, although accurate, are presented in a way that could cause readers to overlook the overall success of the program’s implementation. You review the report one more time and the
findings in the draft regarding operational problems appear to have been reported in a fair and balanced fashion.

**After reviewing the scenario again, please rate how ethically problematic Scenario 1 was.**
1. Very ethically problematic
2. Ethically problematic
3. Slightly ethically problematic
4. Not ethically problematic
5. Not at all ethically problematic

**Please rate how much of an authoritative presence Scenario 1 had.**
1. No authoritative presence
2. Slight authoritative presence
3. Moderate authoritative presence
4. A lot of authoritative presence

**Scenario 2**
In evaluating a school-based youth program which educates students on safe sex at an early age, you have the choice of using either an active-consent or a passive-consent procedure to obtain parental permission. Active consent requires parents to sign and return a form if they wish to give permission for their child to participate in a study. In contrast, passive consent only requires them to sign and return a form if they do not want their child to participate. In general, it is much easier to achieve high participation rates with passive-consent approaches than with active-consent ones. Parents may not allow their children to participate in the survey; which, asks students about their sexual behaviors. Using passive consent will lead to a higher response rate because some parents who oppose the study will simply forget to return the passive-consent form, whereas others who would have opposed the study will fail to read the form in the first place.

**After reviewing the scenario again, please rate how ethically problematic Scenario 2 was.**
1. Very ethically problematic
2. Ethically problematic
3. Slightly ethically problematic
4. Not ethically problematic
5. Not at all ethically problematic

**Please rate how much of an authoritative presence Scenario 2 had.**
1. No authoritative presence
2. Slight authoritative presence
3. Moderate authoritative presence
4. A lot of authoritative presence
Scenario 3
You are evaluating a diversity training program’s outcomes. The program is utilized throughout school districts as a form of teacher professional development in creating a diverse and safe classroom. The evaluation planning runs smoothly and you are even able to collect pre-program data before the training is implemented at the school district. A summer meeting is organized with all the teachers in the school district as a date to collect post-program data. You administer the post-program diversity test and accompanying survey. As you walk around the room, you notice not all the teachers are working on the evaluation item; but rather, they are using this time to socialize and disrupt those who are working. The participant ID numbers are linked to the teachers’ names in a separate excel workbook, which means you have the names of the teachers who do not complete the post-program evaluation items. You are able to identify and report the teachers who did not complete their post-program evaluation items to the program director.

After reviewing the scenario again, please rate how ethically problematic Scenario 3 was.
1. Very ethically problematic
2. Ethically problematic
3. Slightly ethically problematic
4. Not ethically problematic
5. Not at all ethically problematic

Please rate how much of an authoritative presence Scenario 3 had.
1. No authoritative presence
2. Slight authoritative presence
3. Moderate authoritative presence
4. A lot of authoritative presence

Scenario 4
You are conducting a summative evaluation of a community arts program, which is funded through government grants. Key stakeholders include local grant foundations, the community agency who is implementing the program, the local arts guild, the local community center where the program is run, and a few other community organizations that specialize in the arts. You plan to form an evaluation committee out of representatives of the stakeholders so that differing opinions on the evaluation focus and planning can be heard throughout the process. Through the grape-vine, you hear that the local arts guild is known in the community as being against formal assessments of any kind.

After reviewing the scenario again, please rate how ethically problematic Scenario 4 was.
1. Very ethically problematic
2. Ethically problematic
3. Slightly ethically problematic
4. Not ethically problematic
5. Not at all ethically problematic
Please rate how much of an authoritative presence Scenario 4 had.
1. No authoritative presence
2. Slight authoritative presence
3. Moderate authoritative presence
4. A lot of authoritative presence

Demographics

I identify my sex as...
- Male
- Female
- Intersex
- Other (Please specify):
- Prefer not to disclose

I identify my gender as...
- Man
- Woman
- Trans*
- Other
- Prefer not to disclose

What is the highest degree you've completed?
- Bachelors
- Masters
- Ph.D or equivalent (Ed.D., Psy.D., etc.)
- Other

What is/are your degree(s) in?

Which best describes your political orientation?
- Very Conservative
- Conservative
- Moderate
- Liberal
- Very Liberal
How would you describe your current employment status?
- Employed
- Self-employed
- Unemployed
- Student
- Retired
- Unable to work

How many people work for your current employer, counting all locations?
- 1 or self-employed
- 2-9
- 10-24
- 25-99
- 100-499
- 500-999
- 1000-4,999
- 5,000+

What is your occupation and/or professional title?

How familiar are you with the process of gaining approval from the Institutional Review Board (IRB)?
- Not at all familiar
- Slightly familiar
- Modestly familiar
- Very familiar

For how long have you been a member of the American Evaluation Association? (In years)

Is evaluation one of your primary job tasks?
- No
- Yes

For how many years have you been actively practicing evaluation?

In what fields do you conduct evaluation?
Would you consider yourself primarily an internal or external evaluator?

☐ Internal Evaluator
☐ External Evaluator
☐ Other

Please rate how familiar you are with the following works.

Not at all familiar  Slightly familiar  Moderately familiar  Very familiar

The Essential Competencies for Program Evaluators (Stevahn, King, Ghere, & Minnema, 2005)
The Program Evaluation Standards (Yarbrough, Shulha, Hopson, & Caruthers, 2011)

Do you have any additional comments you would like to share?
Appendix C: SJT-E: Authority

SJT-E: Legitimate Authority

INSTRUCTIONS:
• This study is designed to take XX minutes
• For each SITUATION choose ONE from two possible responses. You must only select the ONE response that BEST describes how you would address the situation.
• Answer what you would do as an evaluator practitioner.
• You may sometimes feel you would like more information before answering, but please answer each question based only on the information provided.

1. THE REVISED REPORT SCENARIO
You are working for a philanthropic foundation, evaluating a program that they are a major sponsor for and fund. You, as the evaluator, have recently shared the draft of a final report with the director of the program being evaluated. After reviewing the draft, the program director asks you to tone down one section of the report that describes some operational problems within the program. The director believes that the findings in this section, although accurate, are presented in a way that could cause readers to overlook the overall success of the program’s implementation. You review the report one more time and the findings in the draft regarding operational problems appear to have been reported in a fair and balanced fashion.

Response options:
• You revise the section in question.
• You keep the report as is and disseminate it to the client.

Please provide an explanation for your response to Scenario 1. <text box>

2. THE PASSIVE CONSENT SCENARIO
In evaluating a school-based youth program which educates students on safe sex at an early age, you have the choice of using either an active-consent or a passive-consent procedure to obtain parental permission. Active consent requires parents to sign and return a form if they wish to give permission for their child to participate in a study. In contrast, passive consent only requires them to sign and return a form if they do not want their child to participate. In general, it is much easier to achieve high participation rates with passive-consent approaches than with active-consent ones. Your client, the director of the youth-program, tells you to use passive consent otherwise parents won’t allow their children to participate in the survey; which, asks students about their sexual behaviors. The program director tells you using passive consent will lead to a higher response rate because some parents who oppose the study will simply forget to return the passive-consent form, whereas others who would have opposed the study will fail to read the form in the first place.

Response Options:
• You decide to use the passive consent procedure.
• You decide to use the active consent procedure.

Please provide an explanation for your response to Scenario 2. <text box>
3. THE CONFIDENTIALITY AND PUNISHMENT SCENARIO
The program director of a diversity training program asks you to evaluate their outcomes. The program is utilized through school districts as a form of teacher professional development in creating a diverse and safe classroom. The evaluation planning runs smoothly and you are even able to collect pre-program data before the training is implemented at the school district. The program director organizes a summer meeting with all the teachers in the school district as a date to collect post-program data. You administer the post-program diversity test and accompanying survey. As you walk around the room, you notice not all the teachers are working on the evaluation item; but rather, they are using this time to socialize and disrupt those who are working. The participant ID numbers are linked to the teachers’ names in a separate excel workbook, which means you have the names of the teachers who do not complete the post-program evaluation items. The program director requests you provide the names of the teachers who did not complete the post tests.

Response Options:
- You provide the names of teachers who did not participate in the post-program evaluation items.
- You do not provide the name of the teachers who did not participate in the post-program evaluation items.

Please provide an explanation for your response to Scenario 3. <text box>

4. THE EVALUATION COMMITTEE EXCLUSION SCENARIO
You are conducting a summative evaluation of a community arts program, which is funded through government grants. Key stakeholders include local grant foundations, the community agency who is implementing the program, the local arts guild, the local community center where the program is run, and a few other community organizations that specialize in the arts. You plan to form an evaluation committee out of representatives of the stakeholders so that differing opinions on the evaluation focus and planning can be heard throughout the process. The director of the community agency approaches you after your initial meeting and asks you that when you form the evaluation committee, you leave the local arts guild out because they won’t have meaningful insight to contribute and are known in the community as being against formal assessments of any kind.

Response Options:
- Form the evaluation committee with representation from the local arts guild.
- Form the evaluation committee without representation from the local arts guild.

Please provide an explanation for your response to Scenario 4. <text box>

Rating Scenarios 1-4

Scenario 1
You are working for a philanthropic foundation, evaluating a program that they are a major sponsor for and fund. You, as the evaluator, have recently shared the draft of a final report with
the director of the program being evaluated. After reviewing the draft, the program director asks you to tone down one section of the report that describes some operational problems within the program. The director believes that the findings in this section, although accurate, are presented in a way that could cause readers to overlook the overall success of the program’s implementation. You review the report one more time and the findings in the draft regarding operational problems appear to have been reported in a fair and balanced fashion.

After reading Scenario 1 again, please rate how ethically problematic Scenario 1 was.

6. Very ethically problematic
7. Ethically problematic
8. Slightly ethically problematic
9. Not ethically problematic
10. Not at all ethically problematic

Please rate how much of an authoritative presence Scenario 1 had.

5. No authoritative presence
6. Slight authoritative presence
7. Moderate authoritative presence
8. A lot of authoritative presence

Scenario 2
In evaluating a school-based youth program which educates students on safe sex at an early age, you have the choice of using either an active-consent or a passive-consent procedure to obtain parental permission. Active consent requires parents to sign and return a form if they wish to give permission for their child to participate in a study. In contrast, passive consent only requires them to sign and return a form if they do not want their child to participate. In general, it is much easier to achieve high participation rates with passive-consent approaches than with active-consent ones. Your client, the director of the youth-program, tells you to use passive consent otherwise parents won’t allow their children to participate in the survey; which, asks students about their sexual behaviors. The program director tells you using passive consent will lead to a higher response rate because some parents who oppose the study will simply forget to return the passive-consent form, whereas others who would have opposed the study will fail to read the form in the first place.

After reviewing Scenario 2 again, please rate how ethically problematic Scenario 2 was.

1. Very ethically problematic
2. Ethically problematic
3. Slightly ethically problematic
4. Not ethically problematic
5. Not at all ethically problematic
Please rate how much of an authoritative presence Scenario 2 had.
1. No authoritative presence
2. Slight authoritative presence
3. Moderate authoritative presence
4. A lot of authoritative presence

Scenario 3
The program director of a diversity training program asks you to evaluate their outcomes. The program is utilized through school districts as a form of teacher professional development in creating a diverse and safe classroom. The evaluation planning runs smoothly and you are even able to collect pre-program data before the training is implemented at the school district. The program director organizes a summer meeting with all the teachers in the school district as a date to collect post-program data. You administer the post-program diversity test and accompanying survey. As you walk around the room, you notice not all the teachers are working on the evaluation item; but rather, they are using this time to socialize and disrupt those who are working. The participant ID numbers are linked to the teachers’ names in a separate excel workbook, which means you have the names of the teachers who do not complete the post-program evaluation items. The program director requests you provide the names of the teachers who did not complete the post tests.

After reviewing Scenario 3 again, please rate how ethically problematic Scenario 3 was.
1. Very ethically problematic
2. Ethically problematic
3. Slightly ethically problematic
4. Not ethically problematic
5. Not at all ethically problematic

Please rate how much of an authoritative presence Scenario 3 had.
1. No authoritative presence
2. Slight authoritative presence
3. Moderate authoritative presence
4. A lot of authoritative presence

Scenario 4
You are conducting a summative evaluation of a community arts program, which is funded through government grants. Key stakeholders include local grant foundations, the community agency who is implementing the program, the local arts guild, the local community center where the program is run, and a few other community organizations that specialize in the arts. You plan to form an evaluation committee out of representatives of the stakeholders so that differing opinions on the evaluation focus and planning can be heard throughout the process. The director of the community agency approaches you after your initial meeting and asks you that when you form the evaluation committee, you leave the local arts guild out because they won’t have meaningful insight to contribute and are known in the community as being against formal assessments of any kind.
After reviewing Scenario 4 again, please rate how ethically problematic Scenario 4 was.
   6. Very ethically problematic
   7. Ethically problematic
   8. Slightly ethically problematic
   9. Not ethically problematic
  10. Not at all ethically problematic

Please rate how much of an authoritative presence Scenario 4 had.
   1. No authoritative presence
   2. Slight authoritative presence
   3. Moderate authoritative presence
   4. A lot of authoritative presence

Demographics

I identify my sex as...
   ☐ Male
   ☐ Female
   ☐ Intersex
   ☐ Other (Please specify):
   ☐ Prefer not to disclose

I identify my gender as...
   ☐ Man
   ☐ Woman
   ☐ Trans*
   ☐ Other
   ☐ Prefer not to disclose

What is the highest degree you've completed?
   ☐ Bachelors
   ☐ Masters
   ☐ Ph.D or equivalent (Ed.D., Psy.D., etc.)
   ☐ Other

What is//are your degree(s) in?

Which best describes your political orientation?
   ☐ Very Conservative
☐ Conservative
☐ Moderate
☐ Liberal
☐ Very Liberal

How would you describe your current employment status?
☐ Employed
☐ Self-employed
☐ Unemployed
☐ Student
☐ Retired
☐ Unable to work

How many people work for your current employer, counting all locations?
☐ 1 or self-employed
☐ 2-9
☐ 10-24
☐ 25-99
☐ 100-499
☐ 500-999
☐ 1000-4,999
☐ 5,000+

What is your occupation and/or professional title?

How familiar are you with the process of gaining approval from the Institutional Review Board (IRB)?
☐ Not at all familiar
☐ Slightly familiar
☐ Moderately familiar
☐ Very familiar

For how long have you been a member of the American Evaluation Association? (In years)

Is evaluation one of your primary job tasks?
☐ No
☐ Yes
For how many years have you been actively practicing evaluation?

In what fields do you conduct evaluation?

Would you consider yourself primarily an internal or external evaluator?
- [ ] Internal Evaluator
- [ ] External Evaluator
- [ ] Other

Please rate how familiar you are with the following works.

<table>
<thead>
<tr>
<th>Not at all familiar</th>
<th>Slightly familiar</th>
<th>Moderately familiar</th>
<th>Very familiar</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Essential Competencies for Program Evaluators (Stevahn, King, Ghere, &amp; Minnema, 2005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Program Evaluation Standards (Yarbrough, Shulha, Hopson, &amp; Caruthers, 2011)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Do you have any additional comments you would like to share?
Appendix D: Debriefing

Thank you for your responses; your participation is greatly appreciated. This page describes the background and purpose of this investigation in more detail. **This is an ongoing project so please do not share this information with others who may participate in the future.**

Although the cover story of this study explained that the purpose was to investigate the decisions made in response to potentially problematic scenarios, the primary purpose of this study was to investigate the ethical decisions. Specifically, the decisions made when faced with an ethically problematic situation where either an authoritative figure (client, major stakeholder, supervisor or manager, etc.) makes an unethical request or the situation is vague on how to respond. For the purpose of the study, two form of the situational judgement test was created. One form has an authoritative presence the other does not. Results of both situational judgement tests will be analyzed for differing responses.

There were not obvious right or wrong answers to the scenarios. Ethical dilemmas can be resolved in a number of ways, and still be considered an ethical resolution. Many ethical dilemmas go unnoticed as well. For this reason, you were also asked to rate how ethically problematic each scenario was. These responses will be taken into consideration during analysis and interpretation.

If you have any additional questions about the study or if you would like to receive a copy of the results from this study when they become available, please contact Julia Lamping or Dr. Smith.

As a reminder, this study has been approved by University of Wisconsin-Stouts’ Institutional Review Board (IRB). The IRB has determined that this study meets the ethical standards required by federal law and University policies. This study was also reviewed and approved by the American Evaluation Association (AEA). Any questions or concerns related to the study should be directed towards the thesis student, Julia Lamping, or the assisting professor, Dr. Smith. Any questions or concerns regarding your rights as a participant, contact the IRB Chair, Nels Paulson.

<table>
<thead>
<tr>
<th>IRB Chair</th>
<th>Thesis Student</th>
<th>Assisting Professor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nels Paulson</td>
<td>Julia Lamping</td>
<td>Dr. Tiffany Smith</td>
</tr>
<tr>
<td><a href="mailto:paulsonne@uwstout.edu">paulsonne@uwstout.edu</a></td>
<td><a href="mailto:lampingj0825@my.uwstout.edu">lampingj0825@my.uwstout.edu</a></td>
<td><a href="mailto:smithtif@uwstout.edu">smithtif@uwstout.edu</a></td>
</tr>
<tr>
<td></td>
<td>262.751.8211</td>
<td>(715) 232-2647</td>
</tr>
</tbody>
</table>

By clicking forward you are leaving the study and will not be able to go back to this page.
Appendix E: Non-Authority SJT-E Form Rationale for Scenario One

In Scenario One responses on the Non-Authority SJT-E form, the major rationale was that the report was already “Fair and Balanced” (n = 14) and that the results show areas for program “Improvement” (n = 9). Participants, who provided rationale for revising the report, comments ranged from adding in a few sentences or changing report format to toning down the section in question or revising an alternative section of the report.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does not Revise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair and Balanced</td>
<td>because the report was fair and balanced be begin, it shouldn’t be revised</td>
<td>14</td>
<td>• The operational problems were reported in a fair and balanced fashion.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• I've reviewed the draft and found it to be fair and balanced, so I don't see any need to revise it.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Because the last sentence indicates that the findings were presented in a fair and balanced manner, then it sounds like this report is accurate and needs to be shared as is with the client.</td>
</tr>
<tr>
<td>Improvement</td>
<td>Keeping the operational problems section as is will help the program improve; program improvement hinges on reporting positive and negative aspects; tell it like it is</td>
<td>9</td>
<td>• I was once told that if people are occasionally upset with you, you're doing your job right as an evaluator. Sometimes (although hard) it sheds light on areas that those agencies haven't thought to put a focus.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• I was once told that if people are occasionally upset with you, you're doing your job right as an evaluator. Sometimes (although hard) it sheds light on areas that those agencies haven't thought to put a focus.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• A major purpose of evaluation is to support program improvement. The &quot;operational problems&quot; described should help program staff improve those operations.</td>
</tr>
<tr>
<td>Ethics</td>
<td>comments on how the revision would be a violation of ethics, standards, or competencies</td>
<td>3</td>
<td>• I don't believe this is a violation of the Program Evaluation standards by JCSEE or the evaluation standards given by AERA, APA, &amp; NCME.</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revises Report</td>
<td></td>
<td></td>
<td>- To be ethical, all findings must be reported accurately</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- I think revising the section in question would be unethical.</td>
</tr>
<tr>
<td>Changing Context</td>
<td>Participants suggest changing the interpretation of the section is acceptable as long as data isn't changed and the information presented is still accurate.</td>
<td>22</td>
<td>- Although I would not change my report of the findings, I would consider reordering the presentation of the findings, or adding text that highlights of emphasizes the program's overall success.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- It also may be beneficial to couch problems within positive aspects of the program, while still remaining as fair and balanced as possible in the writing.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- If possible, I would revise the section to make it more clear that the operational problems, while significant, do not affect the success of the organization. As long as the revision does not remove or obscure the problems, I have no issue with rewriting for clarity and purpose.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- It evaluators often have to walk the line between keeping true to the data but also presenting them in a way that are fair and balanced. It does no one any good to overstate data in any direction and in fact does harm to the profession and the project.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- If a section of the report on the operational issues overshadows the overall success of program implementation it needs to be revised to allow for overall balances report.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- The evaluation report should not be misleading overall, and so if the section on operational problems is overshadowing the larger finding, which is that the program is overall successful in its implementation, I would want to revise so that the report is more accurate in the story it tells about the program.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- I would report all of the relevant findings but take the opportunity to revise an alternative section of the report.</td>
</tr>
<tr>
<td>Revises Alternative Section</td>
<td>Participants suggest revising an alternative section of the report.</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Theme</td>
<td>Definition</td>
<td>Count</td>
<td>Examples</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------------</td>
<td>-------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Clients | Participants suggest clients should assist in writing the report and/or disseminating the information to foundations | 8     | • I'm not sure whether I would revise the section or just restructure the report in such a way as to ensure that it was clear that the overall success was positive - I like to lead with the positive so it could be that the placement of the operational findings in the overall report is the issue.  
• I would be open to revising based on my review the format of the report and the executive summary to ensure that the findings are all described accurately. If not, revise. If so, the report stands as written.  
• I would present the report to the client as draft and engage in clarification and discussion of the issues raised. Often, a more detailed explanation of the context of findings is sufficient to reach common ground with clients.  
• The client should have the opportunity to review and provide feedback and discussion if needed before finalizing.  
• Client should now your concerns especially one that may require further explanation. They have to explain it to other stakeholders as well. |
| Other   |                                                                             | 20    |                                                                                                                                                                                                        |
Appendix F: Authority SJT-E Form Rationale for Scenario One

For participants who received the Authority SJT-E form, rationale for keeping the report as is were “Fair and Balanced” (n = 18) and that the “Role” of the evaluator is to provide neutral feedback (n = 9). Of those that rationalized revisions in the Authority SJT-E form, many commented on the “Director” of the program (n = 26) or “Changing the Context” or tone of the section (n = 21).

<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does not Revise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair and Balanced</td>
<td>because the report was fair and balanced be begin, it shouldn’t be revised</td>
<td>18</td>
<td>• If it is fair and balanced, I will keep the report as is. /</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• if the problems were presented fairly I would keep them but use the executive summary to emphasize that the overall program is a success.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Everyone agrees the findings are accurate. It is important that I provide the clearest analysis possible to my primary client, the foundation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• I also don't know what has occurred between the funder and program. it is my job to remain neutral, which it seems the report is. As such, I would not change it at this time.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• As I’m working for the philanthropic foundation, not the program that’s being funded, I’m responsible for providing them with accurate and representative data.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• I think it is our duty as evaluators to report the findings, both good and bad, in a fair way. The data don't lie.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• I would explain to the program director why I see it as fair and balanced and hear his/her reaction to that. I would be curious about what alternate language the director would use and ask for explicit examples.</td>
</tr>
<tr>
<td>Role</td>
<td>It's a part of the role of the evaluator to be neutral and present weaknesses of programs to clients for improvement</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Director</td>
<td>Comments regarding the director of the program</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Theme</td>
<td>Definition</td>
<td>Count</td>
<td>Examples</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------------------------</td>
<td>-------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Ethics        | comments on how the revision would be a violation of ethics  | 6     | • The program may choose to revise the section in some of their literature and reports.  
• I would seriously consider the director's feedback before making that decision, though.  
• Ethical responsibility to present negative findings.  
• I believe it is unethical to deliberately tone down an evaluation.  
• I believe it is unethical to deliberately tone down an evaluation. |
| Improvement   | Keeping the operational problems section as is will help the program improve; program improvement hinges on reporting positive and negative aspects; tell it like it is | 5     | • I would keep all the findings in the report, however I would consider re-organizing the report to highlight successes and offer recommendations for improvement for the operational issues that were found to be problematic.  
• Implementation problems are not incidental to the achievement of results.  
• I have found that when you try to hide things or make them seem better than they really are you are not able to truly make progress or improvements. |
| Revises Report| The director of the program knows more about their program than the evaluator and should be viewed as a subject matter expert. | 26    | • I would have a deeper conversation with the client about that sections, their concerns and their thoughts on how the section should be described.  
• I would have further discussion on the specifics and ask if the foundation has a plan to address the problems and present that plan in the report.  
• Work with the director to report in a fair and balanced fashion.  
• Provide a sentence of 2 directed to data that shows overall success  
• I would revise the language surrounding the findings - but not the findings themselves.  
• There are always ways to paraphrase language to be less "in your face" and if you can do so without altering the integrity of the evidence, it is |
<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Experience          | Comments regarding having been in a similar situation before                | 4     | • This has actually happened to me.  
• This comes up a lot, of course.  
• This has actually happened to me but the PI changed my section in the final report without letting me see it first. |
| Other Revisions     | Other comments regarding revisions to the report                           | 28    | • Depends on the points made. If they have merit will amend.  
• I might include an explanation (note) for the finding based on the rationale offered by the program  
• I'm assuming that the report will be disseminated broadly to audiences not familiar with the operational issues. For the publicly shared report, I would be comfortable toning it down, but not eliminating the negative findings altogether. At the same time, I would recommend to the client that a less varnished report be shared internally, within the foundation. |
| Other               | Other comments not falling into an above theme                             | 6     |                                                                                                                                                              |
Appendix G: Non-Authority SJT-E Form Rationale for Scenario Two

Respecting the parents and guardians rights to informed consent (“Parents & Guardians Respect & Rights”, n = 24) and comments regarding how active consent should always be used (“Active”, n = 15) were major themes in rationale for using active consent in Scenario Two in the Non-Authority SJT-E form.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Active Consent</strong></td>
<td></td>
<td></td>
<td>• Active consent is more respectful toward the values of parents. Too often, researchers disregard the values of parents in the name of their own values or needs of their research. Passive consent is unprofessional, immoral, and unethical for this reason.</td>
</tr>
<tr>
<td>Parental Guardian</td>
<td>Parents and guardians have a right to know about the evaluation, use of active consent to respect the rights of parents/guardians to make decisions for their child</td>
<td>24</td>
<td>• using active consent procedure assures parents are ok with process</td>
</tr>
<tr>
<td><strong>Safe</strong></td>
<td>Consent should always be used; it is unethical to use anything but active consent</td>
<td>15</td>
<td>• Always use active consent with minors.</td>
</tr>
<tr>
<td></td>
<td>To prevent trouble in the future, active consent should be used</td>
<td>12</td>
<td>• Sensitive issues require a higher bar on participant protections</td>
</tr>
<tr>
<td>Depends</td>
<td>If the situation changes, the type of consent procedure may change</td>
<td>9</td>
<td>• The controversy will arise sooner or later. You will be on stronger ethical ground if you deal with it at the outset.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Better to be safe and cover your bases.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• It depends -- active consent seems like it offers highest ethical 'protection' -- but it depends on the school, the age of the students, the educational program</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• For a &quot;hot-button&quot; topic like this, I would do active consent, to make extra sure that parents are on board with student</td>
</tr>
<tr>
<td>Theme</td>
<td>Definition</td>
<td>Count</td>
<td>Examples</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| **IRB**                   | rationale mentioning the IRB or IRB process                                  | 7     | • Required by our IRB  
• I know for a fact that my institution's IRB would not approve passive consent in this instance because some of the contents of the survey are regarding sexual behavior--and I agree with the judgement of my IRB-especially if these students are under 16.  
• I would follow my IRB’s direction - which would most likely be to get consent in a situation like this. |
| **Self-Protection**       | To avoid legal troubles, the evaluator should use the active consent procedure | 4     | • Active consent, because sample size is less important than potential litigation.  
• Sex education is a very controversial subject. I want to protect myself and the school system from any negative ramifications.  
• Given that this requires youth to answer sensitive questions and may have various state and international legal consequences, you should get it. It may be that the states have legal requirements around this. Again, lean on conservative approach with sensitive issues since this may be an issue. Alternatively, if you had funding for legal counsel across various jurisdictions, I would proceed that way before making my choice. |
| **Other**                 |                                                                             | 4     | • If a school system is comfortable with passive consent then I would use it to provide a more representative data |
| Passive Consent           |                                                                             | 4     | • participation. For something less "hot", such as attitudes toward math or writing, I would use passive consent.  
• I generally use active consent whenever I am asking children about potentially sensitive subjects. However, my answer in this case might be influenced by how parents provided consent for their children to participate, as I probably would have tried to ask for consent for the evaluation concurrently with the permission to participate in the program. |
<p>| <strong>Depends</strong>               | if the situation changes, the type of consent procedure may change           | 8     | • If a school system is comfortable with passive consent then I would use it to provide a more representative data |</p>
<table>
<thead>
<tr>
<th><strong>Theme</strong></th>
<th><strong>Definition</strong></th>
<th><strong>Count</strong></th>
<th><strong>Examples</strong></th>
</tr>
</thead>
</table>
| **Response Rate** | Getting high response rate is important for the evaluation | 8 | - I think sexual health programs are a bit tricky and it might depend on the age of the students (if middle school students, I might lean towards active consent but with high school students passive consent).  
- That said, it also would depend on the age of the youth. Since you did not say in the scenario, I might go with the active consent if the youth are very young.  
- So long as adequate consent was gained, I would go with whatever gives me higher participation rates.  
- In general I’d prefer the passive consent procedure as balancing the desire for better and perhaps more unbiased student responses (unbiased in the sense that groups of students are less likely to be systematically excluded from the study).  
- I have been in this situation several times. I go with passive consent since it is much easier to get a good response rate. The cost of getting active consent is very high. |
| **Parents & Guardians** | Parents and guardians are provided more than one change to opt their child out or respond | 5 | - Most parents read the forms and respond actively if they strongly disagree.  
- It is a parent’s responsibility to read notifications sent home from school and proactively respond as required.  
- Rather than requiring active consent, I would be sure that several attempts to reach parents with passive consent forms were made. Perhaps even a full marketing plan to parents with information in a PTA newsletter etc. so there’d be little excuse for them not to know. |
| **Not the Evaluators Problem** | If the school and the IRB allow for passive consent, repercussions are not the evaluator’s problem | 4 | - As long as the IRB at my institution as approved passive consent, I will use it and make it clear to the parents in the passive-consent letter the content of the program.  
- School policy will probably determine which type of consent can be used.  
- It is not my fault the parents do not read the form. |
Appendix H: Authority SJT-Form Rationale for Scenario Two

“Depends” on the situations context (n = 8) and “Response Rate” (n = 8) were major themes for those who chose the passive consent option. Similar to the rationale for Scenario Two Non-Authority SJT-E form, “Parents and Guardians Respect and Rights” (n = 23) and comments regarding use of “Active” consent (n = 21) were major themes for using active consent for Scenario Two in the Authority SJT-E form. Likewise, “Response Rate” (n = 5) and “Parents and Guardians” having adequate opportunity to opt out (n = 4) were major theme for using passive consent in the Authority SJT-E form.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Consent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parental Guardian</td>
<td>Parents and guardians have a right to know about the evaluation, use of active consent to respect the rights of parents/guardians to make decisions for their child</td>
<td>23</td>
<td>• Parents have a right to oppose studies. It is part of mission then, to persuade them that this is an issue of public health.</td>
</tr>
<tr>
<td>Respect &amp; Rights</td>
<td></td>
<td></td>
<td>• Parents must provide consent for non education related data.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Parents have the right to determine what is best for their kids and we as evaluators need to ensure we respect these rules.</td>
</tr>
<tr>
<td>Active</td>
<td>Consent should always be used; it is unethical to use anything but active consent</td>
<td>21</td>
<td>• I would use active consent because of the topic and the age-group.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Because the survey asks about sensitive issues, and children are a protected class, we must use active consent.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Given the sensitive nature of the study and the age of the children, active consent is most appropriate, even though I'm &quot;allowed&quot; to use passive.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• With only this information, I would use the active consent procedure as sex can be a hot topic and there could be lots of push back from parents using the passive consent which could put myself and the program at risk. Other information that would be helpful includes past surveys that have been used and the type of consent, what parents know about the program etc.</td>
</tr>
<tr>
<td>Safe</td>
<td>To prevent trouble in the future, active consent should be used</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Theme</td>
<td>Definition</td>
<td>Count</td>
<td>Examples</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| **IRB** rationale mentioning the IRB or IRB process |                                                                             | 13    | • However, in my personal opinion, the best choice is to choose an active consent procedure. This will likely reduce the potential negative backlash and distrust from parents who feel they were not informed of the study.  
• Given the nature of the study content it would be safer to use the active consent procedure.  
• I would also consult a local IRB board for their feedback on the issue.  
• Given the topic area i believe active consent would be required by the IRB and is the most ethical approach.  
• I believe that an IRB would require active parent consent, especially since the program is with minors and asks about a sensitive topic. |
| **Depends** if the situation changes, the type of consent procedure may change |                                                                             | 7     | • It is too risky in this situation to use passive consent. The only way I would do it is if the school district (superintendent's office) agreed to passive consent.  
• Had the survey been directed at older HS students (such as grades 11 or 12), I may have considered the passive consent procedure.  
• In general, active consent procedures should be used with vulnerable communities. On the other hand, passive consent is likely more appropriate to use among less vulnerable populations, for example, literate adult populations |
| **Ethical Standards** | Mention of evaluation or ethical standards, competencies, or principles | 7     | • I believe that as an evaluator, I have to hold to certain ethical standards.  
• I would probably point to ethical guidelines for evaluators and share my concerns about passive consent not really being consent.  
• Using Passive consent would violate PPRA and would be unethical  
• Again, I would make this decision with ongoing conversations with the client. |
<p>| <strong>Discussion with Client</strong> | More information or discussion is needed before making a final decision with the client | 4     | • I would discuss the pros and cons of using passive versus active consent procedures with the program director. |</p>
<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive Consent</td>
<td>Way of gaining consent that does not require the</td>
<td>5</td>
<td>• I understand that the passive-consent form would be more convenient to gain higher participation but the rational from the program director is unethical. I particularly have issue with the program director's rational on how to make youths of opposing parents participate as it is essentially leading to uninformed consent (e.g. &quot;simply forget to return&quot; and &quot;fail to read the form&quot;).</td>
</tr>
</tbody>
</table>
| Response Rate             | Getting high response rate is important for the | 5     | • Higher participation for a necessary program  
• The scenario states that both forms are available. For that reason, I would use the form that can help net the highest response rates.  
• This is an important program that needs to be fairly evaluated with as little bias as possible. There seems to little risk to the participant so minimizing bias is more important than the risk of harm to the program participant. |
| Parents & Guardians       | Parents and guardians are provided more than one | 4     | • If active consent is not required, I would not use it. Sometimes parents are too overwhelmed to return forms, or don't even receive the forms. Their children should not be denied education because of this.  
• I use the passive consent because the opposite is also true. Many parents who would give consent would not turn in the signed consent. The student still has the option to complete or not complete the survey, or certain questions.  
• I would also ensure appropriate communications are provided to the families and analysis would not include any identifiable information.  
• It depends on the age of the children.  
• Although I say passive-consent, the intended audience for this info is unclear. That my influence my decision. I see this as an opportunity to gather more info, instead of info only from students whose |
<p>| Depends                   | if the situation changes, the type of consent   | 3     | procedure may change                                                                                                                                                                                  |</p>
<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Experience | Rationale is based on previous experience with a similar situation         | 3     | • I believe that if the most important person is the student. As long as they are being treated with dignity, respect and honesty while maintaining their confidentiality, that's what counts.  
• I actually work for an organization that delivers school-based sexual health programs and we use a passive consent form.  
• In our state, we already have passive consent in schools for Healthy Kids Colorado Survey which is for middle and high school.  
• We use passive consent for many sensitive programs and have almost never had a parent complaint. When we've used active consent, we've gotten many frustrated parents who did not have time to fill out the forms, were angry that their children did not participate because they never got the forms, etc., but never had anyone actually opt-out on the basis of content. |
| Allowed    | If it is allowed by the school and/or IRB, passive consent can be used     | 5     | • In the end, I would do what the IRB allowed me to do. I might try to go with passive consent, since I'd want as many kids to participate as possible (and agree that this would lead to more participants). However, I'd need to put this through IRB first. If they were okay with it, I would be too.  
• I would defer to the IRB. Teen sexual rights are a complex matter in the USA and most studies that use active consent are suspect anyway due to the challenge of teen pregnancy being comorbid with both lax parent oversight and those parents who seek to restrict and police their children's sexuality.  
• If the data are confidential and passive consent is allowed in the school's policies, I would use it. |
Appendix I: Non-Authority SJT-E Form Rationale for Scenario Four

Stakeholder participation was the overall rationale for including the arts guild in Scenario Three in the Non-Authority SJT-E form (“All Voices Heard”, n = 30).

<table>
<thead>
<tr>
<th>Theme Name</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Include the Art Guild</td>
<td>All Voices Heard all voices should be heard during the evaluation process</td>
<td>30</td>
<td>• All diverse stakeholder opinion needs to be in the evaluation committee.</td>
</tr>
<tr>
<td>Representatives</td>
<td></td>
<td></td>
<td>• All stakeholders should have a voice</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Let all voices be heard even though you heard someone may be against assessments.</td>
</tr>
<tr>
<td>All Voices Heard</td>
<td></td>
<td></td>
<td>• This is an opportunity to educate the local arts guild about assessments and to understand their perspective. I would be concerned with these questions: Why do they feel this way and how can I engage them in this process?</td>
</tr>
<tr>
<td>Facilitate</td>
<td>Include and facilitate conversation with the guild to see why they are against formal assessment</td>
<td>24</td>
<td>• You should include the local arts guild but have a discussion with all stakeholders before proceeding with the evaluation and allow everyone to express any reservations they might have.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Time to use your negotiation, advocacy, and team building skills!</td>
</tr>
<tr>
<td>Dissent</td>
<td>Negative or counter arguments can improve and strengthen the design</td>
<td>21</td>
<td>• the 'dissenting voice' is often the most informative</td>
</tr>
<tr>
<td>Extend Invite</td>
<td>At least extend the invitation to join the committee, even if they turn it down</td>
<td>12</td>
<td>• Dissent, if managed correctly, can lead to more robust analysis.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Include everyone. Negative opinions frequently allow for improved design.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• I explain and offer representatives from the local arts guild to be part of the evaluation committee, but if they do not want to join, I move on with the evaluation without them.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Chances are, nobody from the local arts guild will want to even participate, but that doesn't mean I wouldn't give them the opportunity.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• I would at least extend an offer. They may offer interesting evaluation efforts to</td>
</tr>
<tr>
<td>Theme Name</td>
<td>Definition</td>
<td>Count</td>
<td>Examples</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Draw Own Conclusions</strong></td>
<td>The evaluator should make their own decision on the invitation and don't rely on hearsay information that could be incorrect</td>
<td>9</td>
<td>- Here say is just that. Draw your own conclusions from your own experience. Should not exclude stakeholders for not sharing a particular perspective.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Do not rely on second-hand information to determine whether to invite the guild. Invite the guild. If guild representatives do not care to participate, they can indicate that.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Also, the grapevine is not necessarily a reliable source of such information.</td>
</tr>
</tbody>
</table>
Appendix J: Authority SJT-E Form Rationale for Scenario Four

“Stakeholder Participation” (n = 48) was the major rationale for including the arts guild in the Authority SJT-E form.

<table>
<thead>
<tr>
<th>Theme Name</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Include the Arts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stakeholder</td>
<td>all voices should be heard during the evaluation process</td>
<td>48</td>
<td>• They are a key stakeholder and must be included, regardless of their outlook towards program assessment.</td>
</tr>
<tr>
<td>Participation</td>
<td></td>
<td></td>
<td>• All of the stakeholders should be included in the evaluation committee if some of them are to be included.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• If they are a key stakeholder, they should have a say if they are included.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• It's not the stakeholder's job to determine whose insight is valuable or agreeable, and taking the word of the community agency director is not a guarantee that their &quot;assessment&quot; of the stakeholder in question is truthful or meaningful.</td>
</tr>
<tr>
<td>Director</td>
<td>The director's job isn't to decide who to exclude or include; facilitate important of art guild's participation to the director</td>
<td>21</td>
<td>• It is not the right of this director to decide who has insights.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• I would find a way to convince the director that it is essential to gather opinions from all groups, and that leaving them out would seriously damage the validity of the study.</td>
</tr>
<tr>
<td>Dissent</td>
<td>Negative or counter arguments can improve and strengthen the design</td>
<td>13</td>
<td>• It may be useful to know what the objections are and find some way incorporating that into the assessment.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Inclusion is important. Also, the project will have more credibility if includes those who disagree or are resistant. Also, it's an opportunity to educate the art guild about evaluation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Little explanation is needed here. Given the comment, the art guild probably has the most informative information on how to improve the program and should therefore definitely be included on the evaluation committee. The evaluator should emphasize with the other members that differing opinions ultimately</td>
</tr>
<tr>
<td>Theme Name</td>
<td>Definition</td>
<td>Count</td>
<td>Examples</td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Facilitate         | Include and facilitate conversation with the guild to see why they are against formal assessment | 9     | • I would work more toward finding a person in the arts guild who is interested in participating. I wouldn't leave out stakeholders based on who is in/out of the directors good graces.  
• It is important to include local arts guild, because they are key stakeholders, and it should be up to the local arts guild to decide whether they would like to be part of the committee or not. It is important to find out why the local arts guild is against formal assessments.  
• However, I would be very intentional in developing a good relationship with the arts guild in order to best understand any concerns they raise. Developing trust and open communication would be very important. |
| Other              | Other comments not falling into an above theme                              | 4     | • Since there are other community organizations that specialize in the arts, you can substitute them for one of your primary stakeholders rather than the local arts guild, because it is a representative a group of people who will be discussing the topic at hand, a community arts program.  
• Difficult to believe that an org not believing in evaluation would participate if invited, so perhaps invite and expect refusal...or ask program is there a different way to include their perspective like an interview around findings versus an eval advisory committee participation  
• My response will depend on the time available. If it is true that the arts guild is opposed to formal assessment, I think they will bog down the process and cause others to reduce their participation. If there is a lot of time and all stakeholders are patient and committed to educating |
<table>
<thead>
<tr>
<th>Theme Name</th>
<th>Definition</th>
<th>Count</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>and bringing along the arts guild, I would insist on including them in the planning because that way I could get their buy-in. In either case, I would still include the arts guild in the data gathering and discussion of results. I would work with them separately to get their input on the process and incorporate it into my method. For me, stakeholder participation in design of the evaluation does not necessarily require a group-based process.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>