WITNESS SELF-EFFICACY: DEVELOPMENT AND VALIDATION OF THE CONSTRUCT

by

ROBERT J. CRAMER

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Psychology
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2009

ABSTRACT

The present study used a Self-Efficacy Theory (SET; Bandura, 1977, 1986) framework to address the need for an outcome measure in witness preparation training. The construct of Witness Self-Efficacy, broadly defined as a witness's perceived ability to testify in court, was developed drawing largely on existing self-efficacy literature. The goal of the study was to establish initial psychometric properties for the Witness Self-Efficacy Scale (WSES).

Participants acting either as mock witnesses or jurors took part in a two-phase study. First, 41 mock witnesses were asked to provide a written account of being falsely accused of an act. They were then briefly deposed about the account, and completed the WSES and construct validity questionnaires. Mock witnesses finally testified under cross-examination from a mock attorney about the accusations. In the second portion, a total of 290 mock jurors (six to eight per group) observed videotaped testimonies of participants from phase one. Mock jurors then provided predictive validity ratings of witness credibility, believability, innocence likelihood, efficacy, and agreement with testimony.

Bi-variate correlation analyses showed that the WSES was significantly positively related with general self-efficacy, social self-efficacy, and general self-confidence. The WSES was significantly negatively related to introversion. The scale displayed non-significant relations with social desirability, depression, self-esteem, and innocence expectancy. Multivariate regression analyses showed that the scale failed to predict any of the primary mock juror ratings. Follow-up multivariate regression analyses showed that mock juror ratings of witness confidence

and the WSES predicted dependent measures, although these independent variables were collinear. Finally, there was an interaction between self and observer ratings of the WSES as it predicted witness credibility.

Results are discussed with regard to theoretical and applied implications. Contrary to SET, self-reported WSE appears to be largely correlated with confidence and fails to predict outcomes on its own. Concerning witness preparation, the WSES possesses solid construct validity and reliability, but requires further testing within a witness preparation training model to assess its practical utility. A self-efficacy based method of witness preparation is proposed as a manner of continued research using the WSES.

ACKNOWLEDGMENTS

I am certain that the following list is incomplete, and words do not to justice to my gratitude for everyone's support through this process. That said, I would like thank my parents from the bottom of my heart for having me (good call on that one), raising me to be such a fine, upstanding citizen (sometimes), and teaching me all I know about education, humor, and fine beverages. Seriously, your unconditional love and support have been a stabilizing presence I could not do without.

To Dr. Stan Brodsky, my advisor, mentor and friend, I cannot express in words all you have done for me over the last five years. I came to Alabama because of you and you have far exceeded my expectations. Your numerous lessons about ant hills, flower buds, Alabama produce stands, watching the squirrels, and God only knows what else have been invaluable. Oh yeah, there is also all that important stuff you taught me about being an open and supportive clinician, compassionate friend, passionate teacher, collaborative researcher, and balanced person. Much of my happiness is to your credit.

I would like to thank Dr. Jamie DeCoster, my teacher, committee member and friend. Your tireless efforts never go unnoticed. I am eternally grateful for the countless hours you have spent supporting me through this and other research endeavors, as well as the reliable laughter and lessons about Bugs Bunny and homoscedasticity. I mean, really, nobody else could ever have taught me such things and actually held my interest.

To Drs. Rebecca Allen and Dave Boles, I would like to thank you for your insightful thoughts, consistent support, and flexibility while serving on my committee. You have both provided fantastic contributions and helped me tremendously in pursuit of this project.

I would like to thank Joel Sogol for your support during this project and in other endeavors over the last few years. I truly appreciate your help with legal issues related to my research, management of my less than stellar moments developing this study, and your unique way with words (of course).

To Drs. Michael Griffin and Shalene Kirkley, although you may not realize it, you both have done as much or more for me in my time in Alabama than anyone. For your roles as mini-mentors of sorts, I cannot thank you enough. I looked to you both for guidance through the program and dissertation, and you always came through. There may have been a conference evening or two that were definitely memorable as well, but you know what they say, what happens in San Diego/New Orleans/St. Pete's/San Fran stays there, right?

To Dr. Marc Boccaccini of Sam Houston State University, thank you for your consultation on research design and scale development, as well as advice undertaking the dissertation process. The doodle during our APLS symposium also helped! Your wisdom is forever appreciated.

I want to thank Dr. Steve Prentice-Dunn for being a consistent resource and friend during my personal and professional development. Your guidance in research design, teaching, and navigating the maze of graduate school has always been appreciated. Your feedback in the initial stages of this project was particularly helpful in keeping me motivated and on track.

To the Graduate Student Association and American Academy of Forensic Psychology, thanks for the cash! I like money.

To Tess M.S. Neal (please note: I distinguished you from other “Tess Neals” you may find on Google, Facebook or other interweb-type sites by adding the proper middle initials☺), your friendship, flexibility, and tireless efforts in the midst of a hellacious graduate school schedule are much appreciated. Your efforts were instrumental to the completion of this project and my degree. I could not have pulled this bad boy off without you! Thank you so much.

I want to recognize my fellow lab members, Jennifer Kelly and Mitch Ziemke for their assistance with the technological aspects of the project. Without your assistance, it likely would have taken me many hours of frustration and several years off my life to get through data collection.

Many friends have lent support by consulting on project issues, sitting through proposal and presentation rehearsals, listening to venting of dissertation-induced procrastination and frustration, keeping me motivated, and helping me enjoy the graduate school experience. This includes, in no particular order, Joe Chandler, Tonya Johnson Chandler, Michelle Hilgeman, Shane Pitts, Sarah Miller, James Soeffing, Erin Gorter, Laura Young, Amie Williams, Brady Duncan, Philip Neal, Tess Neal, Desiree Adams, Ross Grimes, Sara Debus, Emily Wakeman, Wes Webber, and Brian Oliu. I know I definitely do not say this enough, but thank you guys for all the incredible memories over the last five years in the great state of Alabama.

To Emily Patty, my best friend and confidant, thank you. Be it assisting with data collection or keeping me grounded, everything you have done for me over the past year or so means the world to me. I cannot put into words the difference you make in my life.

Enough of this Hallmark stuff. Let’s get this show on the road!

LIST OF ABBREVIATIONS AND SYMBOLS

A	Cronbach's index of internal consistency
F	Fisher's *F* ratio statistic
M	Mean: the sum of a set of measurements divided by the number of measurements in the set
n	Sample size
p	Probability associated with the occurrence under the null hypothesis of a value as extreme as or more extreme than the observed value
r	Pearson product-moment correlation
R^2	A measure of effect size commonly used in regression analysis
SD	Standard Deviation: the variability in a given set of values
t	Computed value of a *t* test
η^2	Eta squared: a measure of the magnitude of a relation between two variables
λ	Lamda: probability statistic used in multivariate analyses
$<$	Less than
$=$	Equal to

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Dissertation Summary

The present study sought to develop the construct of Witness Self-Efficacy (WSE) from a Self-Efficacy theoretical framework. The goal was undertaken in two ways. First, the theoretical and research foundations of WSE were reviewed as a basis for defining the construct. Next, the Witness Self-Efficacy Scale (WSES), a thirty-item questionnaire assessing a witness's perceived ability to testify, was examined for convergent, divergent and predictive validity. Convergent and divergent data were obtained through taped mock witness's testimony and mock witness's completion of construct-related questionnaires. Predictive data were assessed using mock juror ratings of witnesses.

Statement of the Problem

Witness testimony is one of the most utilized (Bernstein & Hartsell, 2005) and influential (Brodsky, 2004) forms of evidence. Witness preparation arose as a common practice for trial consultants and attorneys to bolster a witness's ability to testify (Boccaccini, 2002; Posey & Wrightsman, 2005). Neal (2009) made an important clarification regarding a witness's ability to testify. She noted that the goal of witness preparation should be aimed at aiding witnesses in presenting a clear, effective message. As cited in Neal (2009), the American Bar Association (2001) clearly maintains that witness preparation cannot include strategies to mislead the court. As a consequence of testimony and preparation usage, psychological literature proliferated on witness-related research. For instance, Loftus and colleagues developed

a line of research examining the impact of false information on eyewitness accuracy and confidence, a term that has come to be known as the “misinformation effect” (e.g., Braun & Loftus, 1998; Loftus, 2005).

Germane to efficacious witness testimony, there is preliminary evidence on the effectiveness of witness preparation (e.g., Boccaccini, Brodsky, & Gordon, 2005; Boccaccini, Gordon, & Brodsky, 2004) and an empirically established objective measure of witness credibility (Griffin et al., 2005). Griffin and colleagues (2005) undertook one of the first efforts to develop a theoretically and empirically-supported outcome measure of witness testimony. The Witness Credibility Scale conceptualizes witness credibility as a four-factor construct consisting of knowledge, confidence, trustworthiness, and likeability. This measure has shown utility, especially in expert witness research (e.g., Cramer et al, 2009). What is lacking, however, is a theoretically derived and empirically validated measure of a witness’s ability to testify. Development of such a scale has both theoretical and applied significance.

The present study aims to fill this gap in the psycho-legal literature. The WSES is designed to demonstrate psychometric and conceptual properties towards the goal of assessing a witness’s cognitive, affective, and behavioral perceptions about testifying. Once the WSES is established, these three areas can be assessed in order to identify specific components of testimony to be improved through witness preparation. Self-Efficacy Theory establishes a theoretical backdrop for scale and construct development. Drawing on self-efficacy literature, it becomes possible to define WSE, relate it to similar and disparate constructs, understand the potential implications for juror perceptions, and formulate hypotheses for the present study.

Self-Efficacy: Theory and Application

Social-Cognitive Theory espoused by Albert Bandura (1977, 1986) provides a theoretical foundation for research on perceptions of abilities. As defined by Bandura (1986, 2000), self-efficacy is one's perceived ability to accomplish a behavior in a given situation. In other words, self-efficacy is a cognition or schema about one's performance capability. In a broader perspective, Self-Efficacy Theory is grounded in the empirically supported belief that a person's perceived ability generates or facilitates action and change (Bandura, Barbaranelli, Caprara, & Pastorelli, 2001).

One domain where Bandura and colleagues demonstrated the potency of self-efficacy is child development (e.g., Bandura et al., 2001; Pastorelli et al., 2001). Bandura and colleagues (2001) used a longitudinal design featuring a two-cohort structure to assess cognitive mechanisms impacting children's career plans. Two hundred seventy two Italian children participated across cohorts. Parent and teacher involvement was necessary for objective data analyzing environmental influences on children's career focus. Child participants completed measures of self-efficacy and career interest, and subsequently underwent a major educational shift from junior high school to a choice of 17 professional tracks. Ratings of self-efficacy in this context included self-assessment of one's ability to regulate career interests. Authors were interested in examining how a child's perceived control over career path influenced actual career choices. Additional components of self-efficacy assessed included academic, social, and leisure (extracurricular involvement). Parent and teacher ratings of their own efficacy were also included. Bandura and colleagues (2001) concluded that children's self-efficacy was a direct determining factor of career choice, and mediated the impact of environmental factors such as

parental efficacy in promoting academic success. In sum, children's self-efficacy bore direct and indirect influence on career choice.

Pastorelli and colleagues (2001) investigated the structure of children's self-efficacy across three cultures. Authors noted some differences in self-efficacy beliefs in areas such as academic self-regulatory efficacy, or the ability to control one's own academic efforts and outcomes. Overall, children's social and academic self-efficacies demonstrated solid support and generalizability across cultures. Findings such as these show that one's perceived self-efficacy is a potential contributing agent of change that appears cross-culturally consistent. The principle of self-efficacy as an agent of change has been extrapolated to a variety of contexts such as alteration of diet (Hagler et al., 2007) and boosting teacher effectiveness (Sparks, 1988).

Given the motivational and behavioral impacts of self-efficacy, it is necessary to review the underlying theory on mechanisms and sources of self-efficacy. Bandura (1989, 1993) outlined the basic processes by which self-efficacy is a determinant of our thoughts, feelings and behaviors. First, the author noted that cognitions related to high self-efficacy are high goal setting and increased likelihood to imagine successful scenarios (Bandura, 1993). The opposite is also true; those low in self-efficacy tend to visualize failure. Bandura (1993) noted the cognitive potency of self-efficacy: "a person with the same knowledge or skills may perform poorly, adequately, or extraordinarily depending on fluctuations in self-efficacy thinking" (p. 119). He also noted that degree of self-efficacy is positively associated with effort in information processing and intrinsic motivation.

Self-efficacy operates on an affective level to the extent that it correlates with one's self-esteem (Bandura, 1993). Negative affective coping in the forms of depressive thinking, elevated anxiety, and perceived stress are all associated with low self-efficacy as well (Bandura, 1989). It

appears that self-efficacy is directly and indirectly linked with behavior via cognitive and emotional pathways. Bandura (1993) outlined self-efficacy in students' academic development. At the intrapersonal level, those high in academic self-efficacy have greater success through means including, but not limited to, deeper information processing (e.g., persistent task orientation and test a variety of factors) and increased motivation. In other words, self-efficacy effected positive change in cognitive processing (information processing) and emotional state (desire to succeed in academics).

Crain (2005) summarized Bandura's (1997) four primary ways that one formulates conceptions of self-efficacy. They are actual performance of a behavior, observation of someone performing a behavior, verbal persuasion, and physiological signs. A person arrives at self-perception of efficacy by actual attempts to perform a behavior. Logically, if a person succeeds, his or her self-efficacy increases, and vice versa. In terms of vicarious influences, we judge our self-efficacy in part by others' successes and failures; if we observe a model succeed at a particular task we are more likely to believe we can reproduce such successes. Verbal persuasion, or "pep talks" as Crain said, operates by boosting a person's belief in increased likelihood of success (p. 208). Finally, a person often draws on internal physiological cues to judge his or her level of nervousness, confidence, or competence. One can draw on all four of these mechanisms of self-evaluation in understanding the process by which witnesses formulate self-perceptions of WSE.

The multifaceted manner in which self-efficacy operates allows two assertions. First, measures of self-efficacy apply across various behavioral domains. Secondly, self-efficacy can be quantitatively assessed as an outcome of effectiveness across these domains. Self-Efficacy Theory generated an abundance of general and specific self-efficacy measures in areas including

general self-efficacy (e.g., Chen, Gully, & Eden, 2001; Sherer, Maddux, Mercadante, Prentice-Dunn, Jacobs, & Rogers, 1982), social functioning (Sherer et at., 1982), physical prowess (Ryckman, Robbins, Thornton, & Cantrell, 1982), caregiving (Steffen, McKibbin, Zeiss, Gallagher-Thompson, & Bandura, 2002), teaching ability (e.g., Everett, Price, & Telljohann, 1996), and academic competence (e.g., Yufang, 2004).

Chen and associates (2001) developed a General Self-Efficacy Scale (GSES) that serves as an appropriate methodological comparison for the present investigation. In their first study, authors retained original self-efficacy scale items based on statistical differentiation from Rosenberg Self-Esteem Scale items (Rosenberg, 1965). They also added seven new items to ensure conceptual coverage of the latent construct. Authors labeled this scale the New General Self-Efficacy Scale (NGSES; Chen et al., 2001). Authors used a five-point Likert-type scale for each item, the identical scale utilized in WSES development. Participants for initial NGSES development were 316 upper level psychology undergraduates who completed measures at three time-points throughout the semester (mean time lapse between time periods = 22 days). Authors eliminated six items based on conceptual overlap. To assess content validity, two group panels (one comprised of graduate and one of undergraduate students) compared the NGSES and Self-Efficacy Scale (SES; Sherer et al., 1982) in relation to definitions of self-esteem, general self-efficacy, and other constructs. Panelist results showed that the NGSES possesses superior content validity related to the definition of general self-efficacy then does the SES.

Chen and colleagues' (2001) second study had three goals. First, they sought to distinguish the final version of the NGSES from self-esteem. Second, researchers included predictive analyses to demonstrate the impact of NGSES on actual performance. More specifically, ten work-related tasks were identified because NGSES development was in part

geared toward application in an organizational setting. Finally, authors framed NGSES as a moderator between student exam results and task-specific self-efficacy. They hypothesized that high levels of NGSES would buffer the impact of low academic performance on future perceived success on exams.

A sample of 323 undergraduates provided data at several time-points of self-esteem, general self-efficacy, and perceived competence in the ten professional tasks. Additionally, exam success was measured using students' midterm grades. Overall, the SES showed better test-retest reliability, while the NGSES displayed better internal consistency. Researchers conducted a confirmatory factor analysis using five different models to assess whether general self-efficacy was separate from self-esteem. Authors concluded that the NGSES was more distinct from self-esteem than was the SES. Finally, authors noted that their hypothesized moderator relation "approached significance (p = .06)" (p. 74). In a third study, researchers largely replicated previous findings in a sample of 54 managers in an Israeli MBA program.

Development of the NGSES has particular importance for the present study because it lends insight for design considerations. For instance, in developing a task-specific construct such as WSE, a requisite element of scale development is incorporation of general self-efficacy (GSE). Chen and colleagues (2001) argued against previous scholars (e.g., Bandura, 1997) that general self-efficacy is a more potent predictor of performance outcomes than task-specific self-efficacy measures. Not only is it required to assess the relation between GSE as it relates to WSE, but GSE should be included as a covariate in assessing WSE's ability to predict performance outcomes above other constructs. NGSES development also identifies self-esteem as a crucial content validity relationship requiring attention in order to understand how WSE compares to other types of self-efficacy.

An issue arose in Chen and colleagues' (2001) review of the literature. Differential conceptions of the definition and application of self-efficacy have emerged. There seems to be overall agreement on the definition and specificity of the term "self-efficacy," as proposed by Bandura in various writings. On the contrary, various scholars proffered disparate opinions about the term "general self-efficacy." This distinction has bearing on the present study because it contributes to construct validity relations and analyses. Researchers (e.g., Chen et al., 2001; Judge, Locke, Durham, & Kluger, 1998) have promoted a definition of general self-efficacy as perceived competence across domains. As such, they advocate the stance that general self-efficacy is of more value in predicting direct and indirect effects on beliefs and behavioral performance. This supposition is in direct contrast with those who have argued that task-specific self-efficacies are paramount (e.g., Bandura, 1997). Due to this controversy, general self-efficacy is featured as a prominent construct validity factor in the present study. Comparison of general self-efficacy and WSE would offer further evidence for the superiority of general or task-specific self-efficacy in determining behavior.

Additional self-efficacy scale literature provides valuable information for the present study. This line of inquiry has produced a series of therapist self-efficacy scales such as the Addiction Counseling Self-Efficacy Scale (ACSES; Murdock, Wendler, & Nilsson, 2005), Multicultural Counseling Self-Efficacy Scale-Racial Diversity Form (MCSE-RD; Sheu & Lent, 2007), and the Lesbian, Gay, and Bisexual Affirmative Counseling Self-Efficacy Inventory (LGB-CSI; Dillon & Worthington, 2003). Moreover, several self-efficacy scales were utilized as outcome measures in training programs. For example, Ozer and Bandura (1990) developed and validated scenarios to assess self-defense self-efficacy. Researchers implemented a five-session self-defense training program for women over a five week period aimed at improving perceived

coping skills and beliefs about self-protection. A control group (half of the participants) was wait-listed and completed measures prior to active participation in the self-defense program. A particular strength of this investigation was a six-month follow-up to gather longitudinal data on program effects. Assessment of self-defense self-efficacy facets were matched with specific behavioral components of training, thereby offering direct skills training and modeling to enhance participants' self-efficacy.

The intervention produced two conclusions. First, the self-defense self-efficacy scales displayed solid psychometric data. Second, authors demonstrated a model whereby coping/self-defense efficacy activates two pathways in determining anxiety level and behavior for coping with dangerous situations. Related to the present study, a task-specific self-efficacy showed an impact on cognitive and behavioral outcomes related to that domain. This investigation includes assessment of similar outcomes related to perceptions of ability to testify. Also, subsequent studies utilized the self-defense self-efficacy measure as an indicator of self-defense training effectiveness (Weitlauf, Smith & Cervone, 2000; Weitlauf, Cervone, Smith, & Wright, 2001). The ultimate goal of this investigation was to develop a measure of WSE applicable to witness preparation training in future research.

Summary of Self-Efficacy Theory

Self-efficacy is a pervasive cognitive construct featuring perceptions of performance capabilities. It is a schema that affects emotion and behavior directly and indirectly. People arrive at judgments of self-efficacy through observation, direct performance, verbal persuasion, and physiological cues. An abundance of general and task-specific self-efficacy scales exist in the literature in research domains such as skills training programs, counselor efficacy, and academic performance. While a unified definition of self-efficacy exists, there is some

disagreement about the construct of general self-efficacy. Finally, development and validation of other self-efficacy scales provide valuable insights for the present study.

What is Witness Self-Efficacy?

WSE is one's perceived ability to testify in a clear and effective manner. This includes his or her beliefs about keeping thoughts organized, communicating in a clear and confident way, conveying emotional control, and acting in a professional manner. In this way, the overarching construct of WSE incorporates cognitions, affect, and behavior. However, as empirically demonstrated below, WSE is hypothesized to be a single latent construct. In short, emotional, cognitive and behavioral components are inter-related, as opposed to discrete orthogonal factors of WSE.

To better clarify the nature of WSE, an analogous comparison to Beck and colleagues' (e.g., Beck, 1972; Beck, Steer, & Brown, 1996) description of components of depression is helpful. Depression is conceived of as a single state, or construct. However, Beck and colleagues assert that the construct of depression includes behavioral (e.g., sleep disturbance), cognitive (e.g., suicidal ideation, hopelessness), and affective (e.g., sadness) components. WSE functions much in the same way; there are behavioral (e.g., posture, eye contact), cognitive (e.g., organization of thoughts), and affective (e.g., confidence, nervousness) components related to the underlying construct.

This uni-dimensional conception of WSE is in accord with Bandura's (1997) discussion of general self-efficacy, as well as other narrowly defined types of self-efficacy such as self-defense (Ozer & Bandura, 1990) and social self-efficacy (Sherer et al., 1982).

The Witness Self-Efficacy Scale (WSES) is a thirty-item questionnaire assessing one's perceived ability to testify, possessing both theoretical and applied implications. The WSES

demonstrates good initial psychometric properties. The construct of WSE is hypothesized to be a single latent construct with three facets: Affective, cognitive, and behavioral control on the stand. From a theoretical standpoint, WSE represents an extension of Self-Efficacy theory by testing a global construct in a narrow legal context. Practically speaking, the WSES shows potential utility as an outcome indicator for witness preparation for defendants, experts, and eyewitnesses alike.

Summary of Witness Self-Efficacy Definition

Witness Self-Efficacy is a new conception of task-specific self-efficacy. Drawing on depression literature, an analogy can be made in that both WSE and depression contain cognitive, affective, and behavioral components. Psychometric analyses support a single factor definition of WSE.

WSES Scale Development and Initial Data

Before reviewing the initial data on the WSES, WSES design details are reviewed in comparison to guidelines for constructing self-efficacy scales espoused by Bandura (2005). Bandura provided guidance in formulating items and directions. He counseled scale developers to use the phrase “can” rather than “will” because “can” reflects behavioral perception whereas “will” reflects intent to commit a behavior. This subtle difference is important because self-efficacy is, by definition, a perception or belief about one’s ability. It is our hypothetical judgment of our ability; hence, we should use hypothetical phrasing. The WSES utilizes such terminology in the instructions. Bandura (2005) also suggested including items of varying difficulty in order to tap the entire continuum of perceived self-efficacy. Although no empirical data could be located to define gradations of difficult behaviors on the witness stand, the WSES seeks to adhere to this recommendation because the scale contains items tapping a range of basic

skills (e.g., use of understandable vocabulary, sitting up) to more difficult items to master for witnesses (e.g., remaining calm under cross-examination). The former testifying skills can be accomplished with little effort, whereas the latter may require ample practice or actual testifying experience.

Bandura (2005) also warned of the danger of using items unrelated to the underlying domain to which self-efficacy is applied. In the case of the WSES, items should pertain directly to thoughts, feelings, and behaviors that apply to efficacious testimony. As will be reviewed later, WSES items were derived from anecdotal (e.g., Brodsky, 2004) and empirical (e.g., Boccaccini et al., 2005) facets related to effective testimony. A final piece of Bandura's advice concerning scale items was to pretest and revise items. Initial WSES development featured such revision, both by experts in the field of witness research and by the primary investigator after initial data analysis.

Bandura (2005) also offered suggestions in administration of self-efficacy scales. In order to avoid response bias, the author highlighted the need for privacy in responding in order to prevent socially desirable responding. This can be accomplished in two ways: By allowing the participant to complete the scale alone and using coded numbers to allay concerns that answers would be traced back to that participant. These two methodological considerations have been incorporated into the present design. Coded numbers were used and mock witnesses completed the scale in the Witness Research Lab while the mock attorney and primary investigator were out of the room.

One place where the WSES is different from Bandura's guidelines is the rating scale. Bandura maintains that 100-point scales, or 10-point abbreviated adaptations, were advantageous to smaller scales because respondents fail to endorse extreme responses and smaller scales yield

low reliability. However, Bandura failed to cite empirical evidence in support of these assertions. In fact, scales with narrower ranges consistently yield sufficient reliability as long as items properly reflect the latent construct. The Beck Depression Inventory (Beck, Ward, Mendelson, Mock & Erbaugh, 1961), NEO-PI-R (Costa & McCrae, 1992), and Personality Assessment Inventory (Morey, 1991) are all examples of common measures using smaller response scales that still demonstrate good reliability. The WSES features a five-point response scale because it is deemed large enough to obtain variable responses, but still produce consistent reliability. This notion is supported in the discussion of initial psychometric data.

Finally, Bandura (2005) suggested use of an example item for the purpose of familiarization with instructions. While this is a sensible suggestion, the current form of the WSES does not possess an example item. One was not used because pilot development of the WSES yielded excellent psychometrics. Participants seem to understand instructions provided.

The psychometric data currently available on the WSES elaborate on the definition of WSE. Prior empirical investigations examined only similar outcomes of witness training to witness self-efficacy, namely witness confidence and nervousness (e.g., Boccaccini et al., 2003). These measures, although positively impacted by training, were merely single-item scales. Initial conception and analysis of the WSES (Cramer, Brodsky, & DeCoster, 2006) were geared toward expanding on narrow constructs related to witness testimony. WSES formulation featured item development by six researchers, one professor and four graduate students, in the area of witness research from the Witness Research Lab at The University of Alabama and one professor at Sam Houston State University. The initial 42-item pool is listed in Table 1. The items were developed based upon anecdotally and empirically-supported behaviors and

perceptions related to efficacious testimony (e.g., Boccaccini et al., 2005; Brodsky, 1991, 2004; Cramer, Brodsky, & DeCoster, 2009).

One of the primary sources for WSES item development was research by Cramer, Brodsky, and DeCoster (2009) investigating verbal and non-verbal cues of confidence associated with persuasive testimony. Authors developed three scripts of mock expert testimony based on a Krauss and Sales (2001) article that compared types of expert testimony in capital sentencing cases. The basic content of the testimony, an expert's opinion that the defendant is highly likely to recidivate, was held constant across testimony vignettes. Each script portrayed differing degrees of expert witness confidence (low, medium, and high) based on the following behaviorally defined groupings:

Low Confidence*:* Quivering tone of voice, dysfluencies in speech, vacillating pace of speech, self-corrections, breaks in the flow of words**,** postural awkwardness**,** fixed eye contact**,** saying "you know" to seek assurance, asking for repetition of questions, and signs of anxiety and nervousness.

Medium Confidence: Moderate and stable tone of voice, clarity in speech**,** moderately paced speech, willingness to acknowledge a degree of certainty ("I am reasonably certain"), smooth narrative statements**,** good posture and straight back**,** comfort and poise**,** consistent eye contact**,** hears accurately and responds accordingly.

High Confidence: Loud and strong tone of voice, assertive speech and mannerisms, rapidly paced speech, *always* and *all* statements ("I am certain"), good posture/leaning forward, high fluency of speech.

The behaviorally defined gradations of confidence reflect theoretically or empirically-based cues related to confidence or similar constructs. Therefore, an important distinction for

WSES item development is that these items do not reflect actual witness confidence per se; rather, they denote behavioral groupings for the purpose of assessing varying levels of mock juror agreement and perceptions of credibility. Behaviors associated with the most effective testimony grouping served as a source for WSES items.

Using a between-subjects design, researchers obtained a sample of 317 undergraduate mock jurors who rated two male expert witnesses on the Witness Credibility Scale (Griffin et al., 2005) and likelihood of assigning the death penalty. Overall, results revealed a curvilinear relation between confidence cues and perceptions of credibility, such that the medium level of confidence yielded the highest credibility ratings. This basic finding held significant ramifications for WSES item development because it provided a list of empirically-supported elements of persuasive testimony. Thus, verbal and non-verbal cues associated with medium confidence provided a sound basis for behavioral items included in the WSES. These include an upright posture, consistent eye contact, and willingness to admit a degree of uncertainty in responses.

Additional studies offer support for verbal items of the WSES. For instance, O'Barr (1982) distinguished powerful from powerless speech as they relate to confidence; powerful speech is seen as a function of social status and reflects high confidence, while powerless speech conveys lower social status and low confidence. O'Barr (1982) found that mock jurors rated both male and female witnesses who used powerful, or highly confident, speech as more convincing, truthful, competent, intelligent, and trustworthy than powerless counterparts. Moreover, O'Barr (1982) reported that formal speech (e.g., lay terminology, people's names) was more persuasive when compared to hypercorrect speech styles (e.g., impersonal references such as "the client", technical terminology). Finally, Thomas and McFayden (1995) argued that

people express levels of confidence equal to the degree of assurance in their knowledge or conclusions. In sum, these studies provided a rationale for including WSES items addressing vocabulary use, degree of certainty in statements, and speech style.

A seminal article examining effectual testimony techniques was Boccaccini and colleagues' (2005) work with mock and real criminal defendants. Methodological considerations related to this study are discussed elsewhere, so they will not be reviewed in this section. What is germane to WSES item development is that the researchers delineated a list of 11 testimony delivery behaviors improved by a witness training model. The skills include, but are not limited to, improved posture, decreased fidgeting, sustained eye gaze, and decreased polite speech. Because these skills have been linked with effective testimony and witness training, they are representative of high WSE, and were, therefore, incorporated in the WSES.

Brodsky's work from the perspective of a seasoned expert witness offered more information for WSES item development. For example, Brodsky (2004) advised that the best way to communicate to a jury is to tell a vivid story using easily accessible language and making data connect with a story. In short, he advises to make testimony clear and engaging. An item tapping such ability is included in the WSES. Likewise, within his discussion of the "push-pull", Brodsky (1991) suggested that an expert witness should emphatically agree with blatant gaps or uncertainties in testimony content that the attorney may point out. He recommends such a course of action because the "pull", or the agreement and shifting of one's reply in the direction of the attorney's question, defuses an attack by the attorney and increases perceived credibility. Based on this, and other examples, items assessing appearing credible and admitting uncertainty were included in the WSES.

Brodsky (2004) proffered advice applicable to the emotional components of efficacious

testimony. In his discussion of emotional components of testimony, Brodsky counsels witnesses to lower one's reaction to a posture of emotional serenity. This stance will enable poised, direct responses. His words hold particular truth for coping with aggressive attorney questions. Witnesses are advised to react in an opposite manner by giving quiet, slow-paced answers. Based on these examples, WSES items concerning poise and handling aggressive cross-examination were included as emotional facets of WSE.

Initial Psychometric Properties of the Witness Self-Efficacy Scale

After item development, I collected data for 444 mock witnesses in an online survey format as part of a seminar in scale construction. These data were split into two usable samples, cleaned, and analyzed in a two step format. First, Cronbach's alpha and construct validity data were analyzed (n = 115). The internal consistency of the 42-item scale was equal to .97. Because the scale demonstrated excellent internal consistency, some items were deleted for the sake of size of the scale. However, no single item showed statistically poor relations to the rest of the scale. Therefore, 12 items were eliminated based on conceptual overlap. They were items 3, 6, 7, 12, 15, 16, 22, 27, 31, 32, 38, and 42 from the items in Table 1. Cronbach's alpha for the final 30-item scale is .96. The next iteration of the scale is presented in Appendix A. Construct validity results for the 30-item version are shown in Table 2. The WSES showed a non-significant relation to social desirability, moderate positive correlation with social self-efficacy, and a strong positive correlation with general self-efficacy.

In the second analysis, a confirmatory factor analysis (CFA) was conducted (n = 329) based on a self-efficacy theoretical framework. As highlighted earlier, Self-Efficacy Theory (Bandura, 1997) posits that self-efficacy is comprised of behavioral, emotional, and cognitive facets. With this in mind, the CFA featured a three-factor model, with items divided into factors

representing behavioral, emotional, and cognitive control. The Root Mean Square Error of Approximation (RMSEA) was equal to .083, indicating a non-supported model. A CFA was then run using a one-factor model with all thirty items. Because the RMSEA for the three-factor model was no better than the one-factor model, the WSES remains one total score, thereby reflecting a unitary construct.

Two preliminary assertions can be made from these data regarding WSE as a construct and measure. First, the WSES assesses an internally consistent construct appropriately related to three other concepts. Second, the best interpretation of CFA results to this point shows that WSE is a unitary construct with behavioral, cognitive, and emotional components. Given the current data on WSE, the present study is designed to explore further the nature of both the construct of WSE and the WSES measure. The goal was to develop deeper theoretical understanding of the construct of WSE through WSES psychometric properties and statistical relations with ratings of mock witnesses. To aid in achieving this end, I next review constructs related to WSE.

Summary of WSES Development

The WSES was originally 42 items, but psychometrics and item overlap supported condensing the scale to 30 items. WSES design considerations were largely in accord with advice of Bandura (1997), but dovetailed in the areas of instructions and size of rating scale. WSES items were drawn from empirical research on efficacious testimony and communication styles, as well as anecdotal evidence in the area of expert testimony. Items address affective, cognitive, and behavioral components of testifying.

Construct Validity Relationships

Sechrest (2005) defines construct validity as “the extent to which a measure reflects accurately the variability among objects as they are arrayed on the underlying [latent] continuum

to which the construct refers" (pp. 1584-1585). Put another way, construct validity is the degree to which the construct in question demonstrates appropriate conceptual and psychometric relations with other constructs of interest. In this case, empirical relations between constructs conceptually related to WSE provide confirmation of construct validity for WSE. Preliminary work by Cramer et al. (2006) provides three construct validity outcomes, namely social desirability, general self-efficacy, and social self-efficacy. Expected replication of these findings suggests WSE showing no relation to social desirability, and moderate or strong positive relations with both types of self-efficacy.

Due to the dearth of previous construct relationship investigations with WSE, it is necessary to draw on related forms of self-efficacy to identify construct related measures. From a conceptual standpoint, an element of WSE is that perceived ability to testify includes needing to act in a social situation, or more specifically, testify in front of judges, attorneys, and a jury. Brodsky (1991, 1999, 2004) affords insight to this effect. Testifying requires telling a story to a jury or answering questions. In this way, it functions much in the same way that one's perceived ability in a social setting does. Moreover, Brodsky's various commentaries frame testimony as aimed at teaching the jury. In the case of experts, they must teach jury members through providing scientific data or professional findings. Other witnesses have to persuade by other means. In either case, WSE can be conceptually linked to teaching self-efficacy because both entail persuading or teaching others. Given WSE's link to social and teaching self-efficacy, it is then possible to look at construct-related literature in these two areas to select other measures for validation of the WSES.

Prior validation of teaching self-efficacy scales showed a moderate positive relation with outcome expectancy (Lumpe, Haney, & Czerniak, 2000). Extrapolating from this finding, WSE

would show a moderate positive relation with a witness's perceived likelihood of being found not guilty if the witness claims innocence. Previous investigations of varying social self-efficacy scales resulted in: a) a moderate negative correlation with depression in a sample of youth (Muris, 2001), b) a moderate/strong positive relation with social confidence (Muris, 2001), c) a moderate/strong negative relation with shyness in sample of undergraduates (Smith & Betz, 2000), and d) a positive moderate relation with self-esteem in a sample of undergraduates (Sherer et al., 1982). Drawing on these data, WSE would be negatively related to shyness/introversion and depression, and positively correlated with confidence and self-esteem.

A word of caution is warranted concerning drawing conclusions about constructs related to WSE. When basing these assertions on literature of varying types of self-efficacy, populations and measures, it is necessary to frame them as tenuous. The lack of data currently available renders these hypotheses as tentative; the present study may further our understanding of how WSE is similar or different from other types of self-efficacy.

Summary of Construct Validity Relationships

Limited construct validation of the WSES supports no relation with social desirability, and positive relations with social and general self-efficacy. Because of conceptual overlap between WSE with social and teaching self-efficacies, construct validation measures were drawn from these areas of research. WSE should be negatively related to shyness/introversion and depression, and positively correlated with confidence, self-esteem, and expectancy of innocence.

Witness Self-Efficacy and Confidence

WSE is one's self-perceived ability to perform the act of actually testifying in court. Broadly speaking, WSE is distinguished from confidence in that confidence is only cognitive and affective (e.g., Shrauger & Schohn, 1995); confidence has no tangible behavioral component.

Because WSE is a new concept, there is little direct commentary or evidence differentiating confidence and WSE. However, literature from self-efficacy theory in general can shed light on this distinction.

Confidence reflects a degree of certainty about a perception, event or outcome (e.g., Merkle & Zandt, 2006), while self-efficacy is a specific perception about one's ability to conduct a behavior to a successful degree (Bandura, 1997). Empirical investigations of confidence relate to judgments, events, or outcomes. A common example of this line of research in psychology-law is accuracy of eyewitness testimony. Researchers demonstrated that confidence in one's identification of a defendant does not necessarily imply high accuracy (e.g., Weber & Brewer, 2004; Wells, Ferguson, & Lindsay, 1981). Germane to this discussion is that confidence functions as a degree of certainty about both one's judgment, and, in turn, the outcome of the testimony. Slovenko (1999) offered a similar definition of confidence in the area of expert testimony as the degree of certainty a witness expresses in his or her conclusions. Again, this definition of confidence portrays a relatively broad belief about a person's perceptions of an act or behavior. Self-efficacy, on the other hand, is a specific perceived belief about one's ability to actually carry out a behavior.

Bandura (1997) expounds on the disparities between confidence and self-efficacy. He noted that the term confidence lacks a target of certainty, whereas self-efficacy targets perceived competence in a given behavior. In other words, self-efficacy represents both "affirmation of capability and strength of that belief" while confidence reflects only strength of certainty about an outcome or perception (p.382). Bandura also notes that "confidence" is often employed without a theoretical basis. Self-efficacy, however, is grounded in social-cognitive theory and

considerable empirical data. Likewise, Bandura asserts that confidence is a uni-dimensional trait, but self-efficacy is a multi-faceted belief that shows excellent empirical support.

Theories of Confidence

Bandura's argument that confidence is an overly used term lacking theoretical consistency warrants detailed examination. Also, existing theories of confidence will clarify whether confidence is truly conceptualized as a construct with little or no behavioral linkages. Overall status of confidence theories can then easily be juxtaposed against Self-Efficacy Theory. Extant work on confidence theory pertains to a) studies on the confidence-accuracy relation, b) attempts to define a comprehensive theory, and c) application of social and cognitive theories to confidence research. These areas are reviewed below and followed by conclusions on confidence theory.

Confidence-Accuracy Literature

Much of the confidence literature stems from the relation between confidence and accuracy. For example, Brewer, Sampaio, and Barlow (2005) devised two studies to investigate the "metamemory theory of confidence" (p. 618). In defining their theory, authors proposed a definition in which confidence judgments are based on metacognitive thoughts concerning external confidence cues and subsequent perceptions about their accuracy in recall. Brewer and colleagues identify a "metamemory belief" as a crucial factor in confidence determination because people tend to link highly detailed memory recall with accuracy (p. 619).

In study one, 36 university undergraduates were presented lists of sentences that they would later have to write from memory. Many of the sentences were altered using synonyms for target words (e.g., "hidden" vs. "concealed"), while some were left constant. Two presentation lists were created, each possessing varying forms of some of the same sentences. Participants

were required to write down the finishing portion of the sentence from memory in a booklet, and rate their confidence in recall accuracy. Results supported a positive relation between confidence and accuracy for easy word shifts, but did not hold for complex synonym items. Study two obtained similar results.

These authors explained the latter finding with the notion that participants still rated their confidence high when making numerous errors due to a false belief that full, detailed recall equates to accuracy. In short, participants in high complexity situations employed a "metamemory" judgment of confidence based on depth of recall. Of note is the conclusion that participants drew ratings of their confidence from degree of successful recall. Retrospective judgments of confidence in the metamemory theory discussion support the assertion that confidence is a judgment made after actual attempts at performance of behavior. Applied to testimony, a witness would judge his or her confidence after testifying based on abilities such as fact recall.

Attempts to Define a Comprehensive Confidence Theory

Theories of confidence arguably lack uniformity. Gigerenzer, Hoffrage, and Kleinbolting (1991) provided one of the first empirically-driven attempts at a comprehensive model. Authors describe part of their motivation for deriving the Probabilistic Mental Model (PMM) as an effort to address the lack of consistency in confidence theories. As such, they empirically demonstrate that the theory explains basic confidence-accuracy relations, as well as conditions that mitigate such relations. Moreover, they put forth a new idea: A confidence-frequency principle comparing the tendency to rate confidence in accuracy of one instance versus overall probability of a correct identification. Gigerenzer and colleagues frame the confidence-frequency principle in the context of individual correct answers and overall likelihood of correct answers. Their

results demonstrated that people tend to overestimate their accuracy in single items or questions, but are quite accurate in rating overall accuracy.

Authors identified two underlying cognitive techniques in formation of confidence judgments. The first, a "local mental model (local MM)" represents a quick confidence judgment for a given task based on prior experience and basic cognitive operations (p. 507). On the other hand, we form a PMM if the heuristic judgment is unsuccessful. PMMs frame the given task in the background of all possible instances of a given task. In this way, PMMs evaluate success probability by considering what is necessary in the present task and what is required for success in the given environment. Additionally, PMM formation accounts for extraneous factors other than the target task. The probability of successful witness testimony provides a good example for PMM formation. If a witness forms a PMM, it would compare the specific instance of testifying in that case (e.g., remembering facts of the particular case, dealing with the attorney conducting examination) to overall facets related to the general courtroom environment (e.g., talking to a jury, behavior in a legal setting). A PMM would also take into account other factors (e.g., aggressiveness of cross-examination, unrelated stress in the witness's life at the moment). A final probability rating of confidence would be generated after weighing these factors.

Although Gigerenzer and colleagues reviewed the idea of a local mental model, they failed to explicitly define how it serves as an avenue for shortcut confidence judgments. However, a plethora of research has investigated a Confidence Heuristic Model in which individuals use external cues and past experience to form rapid judgments of confidence (e.g., Price & Stone, 2004; Pulford & Colman, 2005; Yates, 1990; Yates, Price, Lee, & Ramirez, 1996). Price and Stone (2004) defined the confidence heuristic from the standpoint of

overconfidence. They surveyed evidence that people tend to overestimate personal confidence. Also, evidence has begun to accumulate in the literature that perceivers misuse confidence cues to judge accuracy or credibility of a source (e.g., Loftus, 2005; Price & Stone, 2004; Pulford & Colman, 2005). Price and Stone provided 35 undergraduate students scenarios with a reasonably and overly confident financial advisor. Although the sample size was small, a majority (71%) preferred the overconfident advisor. Price and Stone asserted that participants associated high confidence with "competence, knowledge, or correctness" (p. 44). In sum, overconfidence can be associated with perceptions of high credibility or competence.

Shrauger and Schohn (1995) articulated one of the most empirically-derived and comprehensive conventions for understanding confidence. Overall, authors argued for a conception of confidence similar to that of self-efficacy: They proposed the existence of both general and domain-specific confidences. From this perspective, general confidence displays disparate relations to other constructs when compared to confidence in specific tasks. Another conceptual strength of their view on confidence is a distinction between confidence (general judgment of assuredness) and self-worth (judgment of worthiness or esteem).

Shrauger and Schohn developed the Personal Evaluation Inventory (PEI) to assess general and domain-specific confidence. College students were surveyed for life domains of personal importance and items were developed to tap each of these areas. The most frequent domains were: "academic performance, physical appearance, athletics, romantic relationships, social interactions, and speaking before people" (p. 259). Thus, items were developed for these specific areas as well as a constructs reflecting general self-assuredness. The PEI demonstrated good initial internal consistency (alphas ranging from .71-.90) as well as test-retest reliability (.73-.90). The only exception, as expected, was mood state (alpha level equal to .49). Three

studies using college students revealed the following construct validity relations for the PEI. First, there was a high degree of agreement between self and other rated levels of confidence. Second, college students tended to seek out activities they had high confidence in. Hence, the PEI definition of confidence seems to show some predictive association with behavior. Finally, college students with low confidence possess a negative view of their own future as opposed to that of peers.

Shrauger and Schohn (1995) elaborated on their theory of confidence in two important ways. First, they highlighted sources of confidence including judgments from actual performance and stated levels of confidence based on socially desirable responding. Second, they commented on how confidence is a trait detectable by others in social interactions and activities. Others "should be able to judge a person's confidence level" (p. 259). Therefore, Shrauger and Schohn noted how confidence (and confidence ratings) is a product of behavior that feeds into subsequent decisions to engage in a behavior again. The latter part of this conception is similar to the self-efficacy theory position that self-efficacy determines behavioral action and change.

There are apparent flaws in Shrauger and Schohn's perspective. For instance, their scale development and principles are somewhat limited in scope due to the stated focus on college student confidence (i.e., assessing components of confidence that are of significance to members of this population as opposed to a highly generalizable construct). Moreover, their basic definition of confidence raises concern whether they are actually trying to tap a construct too similar to self-efficacy. They defined confidence as, "perceived [assuredness] in competence, skill, or ability" (p. 256). This definition conceptually mirrors that of Bandura's (1997)

conception of self-efficacy outlined previously. Shrauger and Schohn fail to address this overlap. The present study will address this issue.

One confidence theory incorporating self-efficacy was proffered by Stajkovic (2006). Drawing on research from various avenues of psychology and anthropology, the author proposed four domains of employee character that share a common latent bond of confidence. These domains are hope, self-efficacy, optimism, and resilience. Stajkovic (2006) defined confidence as "a certainty about handling something" (p.1209). This definition of degree of certainty is consistent with some others (e.g., Merkle & Zandt, 2006), but lacks specificity of perceived ability in competence or skill offered by Shrauger and Schohn. Stajkovic commented that confidence is an inductive process; we draw conclusions about confidence based on outcomes.

Stajkovic (2006) outlined conceptual parallels between self-efficacy and the other three constructs in order to argue for an overarching confidence core. Self-efficacy is framed as an agent of change similar to hope; empirical data exist demarcating both as contributing agents to behavioral change. Optimism is conceived in the same way. Self-efficacy is emphasized as a basic definitional facet of resilience. In essence, people with high self-efficacy show greater resilience to overcome obstacles. All of these constructs are a product of a combination of a core confidence that enables action, combined with actual skill and desire (motivation) to accomplish a given task. After reviewing 30 motivation theories, the author concluded that a higher-order confidence construct has yet to be adequately considered. Stajkovic offered a differing conception of confidence and self-efficacy worthy of further attention in interpreting the findings of the present study.

Existing Theories Applied to Confidence

Although not a theory of confidence per se, self-perception theory (e.g., Bem, 1967) is one perspective applied to eyewitness confidence. Shaw, Zerr, and Woythaler (2001) posited that people's internal ratings of eyewitness confidence differed from what they were willing to admit publicly. As such, they pointed out a methodological limitation of many confidence rating studies in that only private, confidential paper-pencil methods were used. Shaw and colleagues empirically confirmed that, if participants were exposed to public declaration, self-rated confidence ratings would be lower. These authors explained this notion by maintaining that self-perception theory explicates only internal confidence ratings. On the other hand, "social pressures" decrease our willingness to express high or over confidence in the presence of others (p. 143). Application of self-perception theory to confidence research is a common complicating factor in assessing theoretical relations to confidence. Indeed, many other theories obfuscate this picture including dual process models of information processing (Healy, Light, & Chung, 2005), signal detection theory (Dunn, 2004), and facial recognition theories (Weber & Brewer, 2003).

Conclusions from Existing Confidence Literature

Crucial distinctions arise when attempting to sort through the complexities of confidence theories. While Price's and Stone's (2004) declaration of the existence of a heuristic confidence judgment is in accord with some theories (e.g., Gigerenzer et al., 1991), evidence of overconfidence yielding highest credibility or accuracy contradicts some other findings (e.g., Brewer et al., 2005; Cramer et al., 2009; Sporer, Penrod, Read, & Cutler, 1995). Inconsistent research findings render outlining a uniform theory a problematic task. Several plausible explanations exist for the inability to produce a unifying theory of confidence. For example, the basic definition of confidence appears inconsistent. Behaviorally speaking, Cramer and

colleagues only manipulated a series of verbal (e.g., tone of speech) and non-verbal (e.g., posture) cues hypothesized to reflect varying levels of confidence. From a conceptual standpoint, Shrauger and Schohn (1995) provided a definition similar to self-efficacy in which confidence reflects competence or skill level in a given context. Moreover, they posited confidence as a determinant of behavior as judged from previous outcomes. While definitions such as those by Shrauger and Schohn, as well as Stajkovic, are somewhat general, other definitions such as Slovenko (1999) are narrowly applied and reflect retrospective judgments of certainty that are byproducts of behavior rather than causes.

Confidence research is often influenced by varying physical environments and conceptual contexts such as organizational/business settings (e.g., Price & Stone, 2004; Stajkovic, 2006), the courtroom (e.g., Sporer et al., 1995; Cramer et al., 2009), and college/young adulthood (e.g., Shrauger & Schohn, 1995). Confidence, and subsequently confidence theory, may in fact be a context dependent phenomenon, especially in regard to how judgments of confidence are made and how they impact perceptions of credibility/accuracy. Another contributing factor to a fragmented theoretical basis for confidence is the variety of other theories applied to confidence research (e.g., self-perception theory; Shaw et al., 2001). The proliferation of confidence-accuracy and confidence-credibility research has rendered arrival of a uniform definition of confidence quite difficult.

Summary of Confidence Theories

Confidence theories possess several commonalities, but also present a variety of disparities to be considered in the context of conceptual comparison to self-efficacy. Several theories (e.g., metamemory theory) posit that judgments of confidence clearly emanate from outcomes of performance such as word or sentence recall. On the contrary, other perspectives of

confidence hold that confidence is a determinant of behavior. Moreover, there are mixed findings concerning confidence-accuracy and confidence-credibility, depending on psychological context and theoretical backdrop. One identifiable consensus in the literature is the existence of a confidence heuristic model, even though there are differences in this line of research in relation to accuracy and credibility.

Problematic issues in consistency of confidence theories have been articulated. These include, but are not limited to, inconsistent definitions both in commentary and research, a multitude of research settings, and distantly related theories applied to confidence research. In two instances, there is even conceptual overlap with self-efficacy, either in direct definition or relation to underlying confidence constructs that obscure analysis between confidence and self-efficacy.

Theoretical Value of SET and Confidence

The conceptual difficulties in confidence literature and lack of evidence for WSE have been reviewed. However, what still remains unclear is how the theoretical contribution of the present study can be evaluated. Popper (1957) addressed how the value of a conceptual frame of reference such as SET is ultimately judged. Of note is Popper's philosophical argument that a theory is best validated to the degree it generates a) testable premises or hypotheses, and, as a result, b) empirically supported tenets or positive corroboration across various domains.

Applied to the present study, self-efficacy would demonstrate theoretical value to the degree which Bandura's prior hypotheses are supported in the legal arena. As such, self-efficacy and confidence are discussed in light of Bandura's (1997) assertion that domain-specific constructs possess more predictive value than global traits such as confidence. If this is true,

SET would possess validity and predictive value in outcome variables related to effective testimony above confidence.

One research area where self-efficacy and traits have been compared to task performance is organizational psychology (e.g., Judge, Jackson, Scott, Shaw, & Jackson, 2007). Judge and colleagues examined self-efficacy and Five-Factor Model (FFM) domains of personality (Costa & McCrae, 1992) as predictors of work-related performance using meta-analysis. The FFM domains are emotional stability, extraversion, openness, agreeableness and conscientiousness. Basic correlations showed self-efficacy to be more strongly related to work performance (r = .37) than personality trait domains (*r* ranging from .08-.28). However, in predictive models, self-efficacy offered no significant contribution when added after FFM domains. Mixed results pertaining to self-efficacy's predictive value in this area is complicated by Avery's (2003) findings that self-efficacy predicted ability to voice opinions in a work environment better than four of five FFM domains. Extraversion displayed comparable predictive value.

The example of self-efficacy's functioning related to business outcomes illustrates how SET can be judged in a particular area. Self-efficacy consistently displays positive associations to performance or behavioral outcomes in psycho-organizational research (e.g., Avery, 2003; Bauer, Bodner, Erdogan, Truxillo, & Tucker, 2007; Judge et al., 2007; Stajkovic & Luthans, 1998). However, there are mixed results related to self-efficacy above and beyond general traits. Thus, SET is somewhat validated in this domain. The present study represents the first thorough examination of SET in a legal domain. I sought to assess SET as a predictor of behavior in testifying and compare self-efficacy to global traits, specifically confidence, to assess whether domain-specific attitudes are more potent predictors than traits. Either way, the present study possesses value in that SET was tested in a testifying scenario for the first time.

Witness Self-Efficacy and Witness Confidence

Drawing on the links between confidence and self-efficacy in general, comparisons can be extended to witness confidence and witness self-efficacy. Witness confidence (WC) stems from a positive belief in the nature of a witness's knowledge or observations. In other words, it pertains directly to the degree of certainty in content and statements made on the stand. Furthermore, WC judgments are often made ex post facto; witnesses may justify their confidence level based on the actual outcome of testimony. WSE, on the other hand, is a belief that directly affects performance on the witness stand. While it includes degree of certainty in ability to testify, WSE is more complex than WC because it is a malleable construct that impacts, and, in turn, is impacted by, cognitive and emotional factors related to testifying, as well as the act of testifying. WC does not directly target or influence testifying. Rather, it can be a judgment or perception based upon the outcome of testimony.

These two constructs also differ in that WC does not target specific behaviors, but rather is judged from behavioral indicators (see Cramer et al., 2009). WSE, however, addresses cognitive (e.g., organizing thoughts), emotional (e.g., remaining calm), and behavioral (e.g., consistent eye-contact) facets of testifying in that building WSE fosters improvements in these three areas. WSE also originates from theoretically and empirically-supported conceptions of general, social, and teaching self-efficacy. However, the construct itself has yet to be validated. Literature on confidence offers a more fragmented definitional basis for WC when compared to WSE. WC does have more empirical data, but this information pertains mainly to raters (e.g., mock jurors) judgments of a witness's confidence. Finally, WSE provides practical use for an outcome of witness preparation training after targeting improvement of self-efficacy beliefs and

testimony delivery skills. Because WC is a narrower construct, its practical implications are limited mainly as an indicator of witness credibility research.

Table 3 outlines the differences between self-efficacy and confidence, as well as witness self-efficacy and witness confidence. Given both the conceptual similarities and differences between types of self-efficacy and confidence, it stands to reason that WSE should show a moderate relation to a witness's general feeling of self-confidence.

Summary of WSE and Witness Self-Confidence

WSE is a construct consisting of affect, cognition, and behavior. Further, WSE precedes actual performance of testifying, but lacks sufficient empirical support. WC is a narrowly defined construct resulting from post-testifying judgments about one's own or other's testimony. There is a dearth of empirical data supporting a consistent definition of WC. WC is primarily used in credibility/accuracy research, whereas WSE is a potential outcome measure for witness preparation.

Witness Self-Efficacy and Outcome Indicators

Previous research on predictive relationships with self-efficacy scales showed a moderate relation between time spent teaching with teaching self-efficacy (Brenowitz & Tuttle, 2003; Everett et al., 1996). If this holds for WSE, testifying experience may relate to WSE. Therefore, this was included as an exploratory variable, although it seems unlikely that enough participants would have such experience as to warrant statistical attention. This line of inquiry also shows no relation between social self-efficacy versus others' ratings of actual social skills in a sample of individuals with intellectual disabilities (Payne & Jahoda, 2004). This finding is complicated by conclusions by DuCharme and Bachelor (1993) who found self and other ratings of social competence (self-efficacy) to correlate in maladjusted undergraduates, but that social

competence was overestimated in healthy undergraduates when compared to observer ratings. If these findings are extended to WSE, mock juror ratings of WSE and witness ratings may not correlate in healthy persons. The inconsistent perceptions of WSE render a hypothesis as exploratory. Finally, Sherer and colleagues (1982) found social self-efficacy to predict educational and military success. Therefore, WSE is hypothesized to predict testifying success defined by positive correlations with credibility, perceived innocence, and agreement with the witness.

Defining Credibility

The expected associations between a witness's self-efficacy and both expectancy of innocence and juror agreement are arguably easy to explicate. If a witness expects to testify well, and does so, then he or she will likely generate juror agreement. Agreement may lead to verdict results in the witness's favor. However, the link between WSE and juror ratings of witness credibility warrants more attention, largely due to the complexities in defining credibility.

Origins of witness credibility research are rooted in source credibility discussions within social psychology literature. For example, McCroskey and colleagues (e.g., McCroskey, 1966; McCroskey & Young 1981) established much of the conceptual and empirical groundwork on source credibility. McCroskey (1966) employed the term "ethos" when discussing source credibility. He designed (McCroskey, 1966) and revised (McCroskey & Telen, 1999) a source credibility scale including items such as good will, competence, and trustworthiness. In a 1981 discussion of the state of source credibility theory and research, McCroskey and Young highlighted eight factors that are analytically supported components of credibility: Sociability, size, extroversion, composure, competence, time, weight, and character. Authors concluded that

these eight domains could be collapsed into two overarching subcomponents of credibility: Competence and character. Brodsky (2004) defined similar constructs of knowledge and trustworthiness, respectively, in the context of expert witness credibility.

Cole and McCroskey (2003) reported that credibility was negatively related to verbal aggression and communication apprehension. By extension, credibility would be positively related to high WSE because a person high in WSE is not verbally aggressive or hesitant about communication skills. Witnesses who possess high self-efficacy will convey testimony skills and characteristics that may persuade jurors via both high and low effort routes of persuasion (e.g., Petty & Cacioppo, 1986; Smith & DeCoster, 2000). Persuasion will, in turn, result in mock juror ratings of high components of credibility. As stated earlier, Griffin and colleagues (2005) delineated an empirically-based four-factor structure of witness credibility suitable for research efforts such as the present investigation. The four factors are witness likeability, trustworthiness, knowledge, and confidence. Use of a multi-faceted definition of witness credibility will offer a nuanced depiction of exactly how WSE predicts credibility, thereby providing invaluable construct validation information.

Design Considerations

Mock jury research often necessitates careful balance between external validity and methodological design. The following two sections address justifications for decisions made in the present investigation concerning this conflict.

Justifying Use of Mock Jurors

Numerous concerns exist over the use of mock juror research despite its application to various law-related topics including, but not limited to, comprehensibility of jury instructions (e.g., Rose & Ogloff, 2001) and expert witness testimony (e.g., Krauss & Sales, 2001). Critics of

the use of mock juror pools argue that convenience sampling and constricted demographic characteristics of college-age samples prevent proper generalization of findings from such investigations. However, moderate evidence exists to dispel such concerns.

Bornstein's (1999) review of mock juror research examined ecological validity by comparing trial presentation styles (written vs. videotape) and samples of mock jurors (college students vs. prospective jurors). Among the pertinent findings of Bornstein's review are that student mock juror usage is increasing and that student samples yielded similar verdicts in 21 of 26 studies reviewed. MacCoun's (1989) review of mock jury research produced similar results in that jury-eligible and student samples resulted in nearly identical verdict outcomes. Also germane to the present study is that previous mock juror research showed student and jury samples as equally persuaded by expert witnesses (Hinkle, Smeltzer, Allen, & King, 1983). However, Bornstein (1999) warned researchers involved in jury simulation research that the judicial system has adopted an ardent stance on psycho-legal research. Thus, it is prudent to undertake mock jury research that utilizes simple methods and allows for easy replication. Additional support for use of mock jurors comes from Dunn's (2003) findings that methodological differences (e.g., mock jurors vs. community sample) in jury decision making account for a miniscule amount of variance.

Taken together, the evidence surveyed offers a cogent rationale for the use of mock juries in psycho-legal research. Indeed, Bornstein (1999) inferred that mock juror research aimed toward theoretical ends is easily justifiable. The present project is an example of such a study because conceptual development of witness self-efficacy, especially in relation to confidence and performance outcomes, is a stated goal. Because principles in theoretically-driven mock juror

research are often extended to other settings and populations, just as research using college students is, use of mock jurors in the present study is defensible.

Does Mock Juror and Mock Testimony Research Possess Practical Value?

Although Bornstein's (1999) conclusions offer strong support for the theoretical value of mock juror research, a question of practical utility for attorneys and legal decision makers remains. Critics of mock jury research, and by analogy mock testimony, research can argue that psycho-legal research fails to actually impact legal process and outcomes. On the contrary, Haney (1980) maintained that psychology-law research could yield valuable information to alter the way in which the justice system operates. In a follow-up commentary, Haney (1993) stated that the psychology-law field has made notable contributions to influence the justice system. He cited psychology's part in changing policy in some jurisdictions concerning the way in which eyewitness identification is considered. Other areas where psycho-legal research has affected positive change in the legal arena include clarity of jury instructions and caution applied to predictions of dangerousness in some states (Haney, 1993).

The issues of jury instructions and jury selection are especially pertinent to the present study in justifying the external value of mock juror research. For example, research on jury instructions (e.g., Severance & Loftus, 1982, Severance & Loftus, 1984) used mock jurors (undergraduates) to demonstrate complexities and weaknesses in jury instructions. As cited in Haney (1993), research using mock jurors has been included in legal decisions such as three states' movements to revise jury instructions. The fact remains that jury instruction and other research (see Haney, 1993 for full review) are often used in legal decisions despite limitations on external validity. This principle can be extended to mock witness testimony; although there are

inherent limitations in ecological validity, lab-based research offers at least preliminary findings that can inform witness preparation and legal decisions.

Hypotheses

In light of the convergent and divergent relationships WSE demonstrated so far, and the additional constructs related to general, social, and teaching self-efficacy validation, the primary hypotheses of the present study were:

H1: Self-reported WSE will demonstrate a non-significant relation with social desirability.

H2: Self-reported WSE will demonstrate moderate positive relations with social self-efficacy, innocence expectancy, self-confidence, and self-esteem.

H3: Self-reported WSE will demonstrate a strong positive relation with general self-efficacy

H4: Self-reported WSE will demonstrate moderate negative relations with depression and introversion.

H5: Self-reported WSE will positively predict mock juror ratings of credibility, agreement, believability, and innocence.

H6: Self-reported WSE will predict variance in mock juror ratings above and beyond confidence and general self-efficacy.

Exploratory Hypotheses:

H7: Self-reported WSE will positively predict observer (juror) ratings of WSE.

METHODOLOGY

Participants

Mock witnesses and jurors were drawn from an introductory psychology research pool and received course research credit. Forty-one students served as mock witnesses for the taping portion of the investigation. Three areas of psychological research that can offer support for choice of a relatively small sample size of mock witnesses are witness preparation effectiveness, evaluations of therapy outcome, and self-efficacy scale development. For example, Boccaccini and colleagues (2003) obtained 55 mock defendants to evaluate the effectiveness of witness preparation on several outcome measures. Their design featured four conditions based on training and what attorney conducted witness examination. In a similar vein, treatment outcome studies on cognitive-behavioral interventions for insomnia (e.g., Soeffing et al., in press) and depression (e.g., Backenstrass et al., 2006) feature sample sizes ranging from the low thirties to mid-fifties. Finally, Ozer and Bandura (1990) used a sample of 43 women in assessing pathways related to anxiety and behavior change in the backdrop of a self-defense training program. They developed a self-defense, or coping, self-efficacy scale that assessed one's perceived ability to detect and handle dangerous situations.

Still regarding sample size, initial scale development assessing basic validity relations may only require a small sample. Afterward, the scale can be adapted to more accurately reflect the latent construct. It can then be tested under more externally valid, time intensive conditions such as use of real criminal defendants. In fact, Boccaccini and colleagues' witness preparation research followed this pattern (see Boccaccini et al., 2005).

Participants acting as witnesses had a mean age of 20.37 years (*SD* = 4.00). The sample consisted of 18 males and 23 females. They reported their ethnicity as Caucasian (n = 28), African-American (n = 11), Latin-American (n = 1), and Bi-racial (n = 1). Participants identified religion Atheist/Agnostic (n = 14), Southern Baptist (n = 13), Methodist (n = 6), Protestant (n = 4), Jewish (n = 1), and Buddhist (n = 1). Two participants failed to identify their religion. Four individuals had testified in court prior to this experience and one had previously served on a jury.

Following taping of mock witnesses, each of the 41 mock witness tapes were observed by groups of approximately six to eight mock jurors. Therefore, the mock juror participant total was 290 for the jury decision making portion of the study. I elected to use approximately six mock jurors per group to reflect the externally valid situation of some federal juries. The mean age of this sample was 18.90 years (*SD* = 2.11). The sample consisted of 82 males and 208 females. They reported ethnicity as Caucasian (n = 234), African-American (n = 40), Latin-American (n = 7), Asian-American (n = 3), and Bi-racial (n = 6). Participants identified their religion as Christian (n = 116), Baptist (n = 66), Catholic (n = 51), Methodist (n = 30), Atheist/Agnostic (n = 7), Protestant (n = 2), Jewish (n = 1), and 'other' (n = 17). Fifteen individuals had testified in court prior to this experience and one had previously served on a jury.

Materials

All questionnaires but one are included in Appendix A. For a copy of original copyrighted materials (BDI-II) please consult the authors. Table 4 lists descriptive and psychometric data for each scale based on data obtained in the present study. Cronbach's alpha values were not calculated for single-item scales. Also, one participant failed to complete juror ratings; therefore, the overall sample was 289. The measures were:

Demographics. The demographic form included mock witness/juror age, sex, race, religion, parental occupation, jury duty experience, and testifying experience.

Mock witness measures

Witness Self-Efficacy. Witness Self-Efficacy was measured with the Witness Self-Efficacy Scale (Cramer, Brodsky, & DeCoster, 2006). This 30-item scale assesses one's perceived ability to testify in court (see Appendix A). Each item is a statement pertaining to a person's perceived ability to testify, including statements addressing emotional, cognitive, and behavioral control. Items are rated on a 5-point Likert scale ranging from 1 (*Not well*) to 5 (*Very well*), and summed for a total score. Preliminary data analyses showed an overall Cronbach's alpha of .96 for the scale.

Social Desirability. Social desirability was assessed using the Marlowe-Crowne Social Desirability Scale (MCSDS; Crowne & Marlowe, 1960). The scale consists of 33 true/false items, 15 of which are reverse scored. All positive responses to items are totaled for a composite social desirability score. In a summary of the MCSDS, Paulhus (1991) reported average scores ranging from 13.3 (*SD* = 4.3) to 16.4 (*SD* = 6.5), as well as Cronbach's alpha levels from .73 to .88 across studies.

Social and General Self-Efficacy. Social and general self-efficacy was measured with the Self-Efficacy Scale (SES; Sherer et al., 1982). It consists of 30 items with two subscales: Social self-efficacy (six items) and general self-efficacy (17 items). There are seven filler items. Each statement is rated on a five-point scale ranging from 1 (*Disagree strongly*) to 5 (*Agree strongly*). Fourteen of the items (11 general and three social) are reverse scored. Cramer et al. (2006) reported a Cronbach's alpha of .87 for general self-efficacy and .70 for social self-efficacy.

Witness Innocence Expectancy. Witnesses rated their perceived likelihood of being found innocent using a 10-point Likert item. Higher values reflected an increased probability of being found innocent.

Self-Confidence. Self-confidence was measured using the Personal Evaluation Inventory (PEI; Shrauger & Schohn, 1995). The scale consists of 54 statements, each rated on a four-point scale (*Strongly disagree* to *strongly agree*) with higher values reflecting greater confidence scores. The scale consists of the following eight subscales (alpha reliability in parentheses): General (.71), Speaking (.86), Romantic (.86), Athletics (.90), Social (.82), Appearance (.83), Academic (.77), and Mood (.85) (Shrauger & Schohn, 1995). Only the seven item General Self-Confidence subscale was used for the purpose of gathering validity data for the WSES.

Self-Esteem. Self-esteem was assessed using the Rosenberg Self-Esteem Scale (Rosenberg, 1965). The scale consists of ten statements (e.g., "I certainly feel useless at times"), each rated on a four-point scale ranging from 0 (*Strongly disagree*) to 3 (*Strongly agree*). Five items are reverse scored. All items are totaled for an overall self-esteem score. Cronbach's alpha ranged from .77 to .88 (e.g., Rosenberg, 1986).

Depression. Depression was measured with the Beck Depression Inventory – 2nd edition (BDI-II; Beck, Ward, Mendelson, Mock & Erbaugh, 1961), a 21-question survey assessing cognitive, behavioral, and affective depressive symptoms. Each item contains four statements pertaining to the same symptom, and participants are instructed to select the one that best describes them. Each item is scored zero to three, and all items are aggregated for a total depression score. Internal consistency ranges from .73 to .92 with a mean of .86 (Beck, Steer, & Garbin, 1988). In order to avoid ethical complications involved with inquiring about suicide,

this item was deleted from the BDI. Reliability was calculated to ensure that this item deletion did not impact psychometric properties in the present study (see Table 4).

Introversion. Introversion was examined using the 10-item Introversion/Extroversion scale from the 50-item International Personality Item Pool (Goldberg, 1999; IPIP, 2001). This is a brief measure based on the NEO-PI-R (Costa & McCrae, 1992) that measures five personality domains: neuroticism/emotional stability (N), extroversion (E), openness to experience (O), agreeableness (A), conscientiousness (C). The personality factor measure consists of 50 questions scored on a five-point likert scale. Alpha reliabilities have been reported as follows: .86 for N, .87 for E, .84 for O, .82 for A, and .79 for C (Goldberg, 1999; IPIP 2001). E was the only scale of interest for the present study. Therefore, others were not included in the questionnaire packet. Although higher scores on this scale reflect higher extroversion, all items keyed in the direction of extroversion were reverse-scored for the sake of consistency with hypotheses. Therefore, higher scores denote increased values of introversion.

Mock juror measures

Witness Credibility. Witness credibility was measured using the Witness Credibility Scale (WCS; Griffin et al., 2005). The scale contains twenty items, each rated on a ten-point Likert scale. Factor analyses yielded four separate, robust domains: Confidence, likeability, trustworthiness, and knowledge (Griffin et al., 2005). Alpha coefficients have been reported for each subscale and are as follows: Confidence (.89), likeability (.86), trustworthiness (.93), and knowledge (.86). All four subscales are totaled for an overall credibility score. Alpha for the total score was .95. For the present study, the knowledge domain was not used because it is most applicable to expert witnesses. The WCS was developed based on perceptions of expert

testimony, and items in the knowledge domain (e.g., “scientific-unscientific”) are germane only to expert witnesses.

Juror Rating of Allegation Severity. A ten-point item asking “How severe do you consider the allegation against the defendant?” was included to control for the severity of crime in each particular case of testimony. Each of these ratings ranged from 1 (*harmless*) to 10 (*extremely severe*).

Juror Rating of Guilt and Agreement. Mock jurors rated the likelihood he or she would find the witness “not guilty” with a ten-point Likert item. Higher values reflected an increased probability of being found not guilty. Two similar items concerning mock juror perceptions of agreement with testimony and the witness’s believability were also included. Witness believability and agreement are also rated on ten-point scales as detailed (see Appendix A).

Juror Rating of Witness Efficacy. Mock juror rating of Witness Self-Efficacy was accomplished with the WSES (see scale details above). The only differences were in the instructions and focus of the items. Mock jurors were asked to rate the witness’s performance during testimony on each item. Items using self-references (e.g., “I”, “my”) were altered to reflect assessment of the witness.

Procedure

Validation of the WSES necessitated three steps: 1) training a mock attorney, 2) taping witnesses, and 3) obtaining juror ratings. Attorney training required training of one graduate student from the Witness Research Lab by the researcher and an expert in developing cross-examination questions. There is precedent in witness-related research for using psychology-law graduate students as mock attorneys (e.g., Boccaccini et al., 2003; Cramer et al, 2009). Moreover, a graduate student from the Witness Research Lab, as compared to a trained actor,

possesses familiarity with psycho-legal terminology. While it was equally acceptable to consider a law student, the member of the Witness Research Lab was also familiar with the research setting. Because neither law or psychology-law students possess a wealth of courtroom experience, the psychology-law student was chosen due to the familiarity with setting and terminology.

The following were used in training the mock attorney: HM Revenue & Customer guidelines for cross-examination (HMRC, 2007), a DVD on designing cross-examination questions (Pozner & Dodd, 2006), review of cross-examination trial transcripts in the public domain, and consultation with an expert to direct training and practice for formulation of cross examination questions. The HMRC guidelines offer principles to follow in developing cross-examination questions such as how to best attack witness character and to use leading questions. These guidelines offered advice in developing an externally valid testifying situation because they provide reminders of keeping questions brief, following a logical sequence, including data not brought up in direct testimony, and testing the witness's ability to remain poised and recall case facts. Likewise, the video offers guidance in the use of techniques such as close-ended questions, and monitoring your own and the witness's body language and voice inflection. Pozner and Dodd's (2006) instructional video also bolsters the external validity in that the mock attorney was able to use actual strategies from the courtroom. For example, the DVD takes a trainee through typical sequencing of questions and supplies rules for handling unexpected changes in the course of testimony. Facets of cross-examination aimed at challenging the witness will yield a realistic testifying experience. Trial transcripts offer practical examples of how one proceeds with cross-examination. Additionally, consultation with an expert was deemed appropriate in order to enhance external validity of questions used; the expert was a trial

consultant who has assisted in writing cross-examination questions before. Training consisted of approximately ten hours of individual and collaborative review and ceased once the researcher and expert felt that the mock attorney had demonstrated clear skill in developing cross examination questions.

The witness taping portion of the study was composed of three steps that were completed in one two-hour session per mock witness. First, undergraduate students acting as mock witnesses underwent notification and review of consent procedures both verbally and in written format (see Appendix C). They were then asked about issues of which they have been unjustly accused. They were required to provide a written account of the accusations against them. Table 5 contains information concerning types of accusation made against mock witnesses. As advised by Boccaccini and colleagues (2003), researchers limited the scope of the accusations so as to avoid potential harm. Participants were asked not to provide allegations pertaining to: a) elder or child abuse or neglect, b) incidents containing excessive psychological or physical harm, or c) anything that could lead to a lawsuit or criminal charges. The innocuous nature of allegations, while a methodological limitation, was justifiable because it complies with *APA Ethical Guidelines* (2002) to avoid causing harm to research participants. The exact form that was employed is featured in Appendix D.

Next the mock attorney developed cross-examination questions from the written material provided by the mock witness and follow-up deposition-type probing designed to gather additional relevant information from the witness. The depositions were non-intrusive and relatively brief, often lasting approximately five minutes. While the attorney completed these questions, mock witnesses completed the WSES, demographic form, and construct validity questionnaires. Ethical considerations can arise in instances where participants obtain high

scores on the BDI-II. In order to address this, the primary investigator scored the BDI-II during the session. Participants obtaining a score of 20 or above (i.e., scoring in the moderate to severe range of depressive symptoms; Beck et al., 1996), were encouraged to seek assistance from a mental health professional and were provided contact information for the University of Alabama Psychological Clinic, University of Alabama Counseling Center, and University of Alabama Police Department (see Appendix E). One participant had a score over 20 and was provided referral information.

Finally, witnesses underwent a series of cross-examination questions by the mock attorney while being videotaped. Each taping lasted approximately 4-10 minutes. I elected to use only a cross-examination condition based on previous research. Boccaccini et al. (2003) reported witness preparation to be effective in coping with cross-examination. Cross-examination provides the most stressful, and perhaps realistic, testifying condition in which to assess WSE. As stated earlier, one goal of WSES development was to use it as an outcome indicator for witness preparation. Because witness preparation is effective for cross-examination, it is logical to develop the scale using such a scenario.

The cross-examination questions were designed to elicit details concerning the event of which the witness was accused. As Boccaccini and colleagues (2003) pointed out, it is impossible to anticipate all answers the mock witness would provide. Therefore, the mock attorney was permitted to ask additional questions to those designed based on witness written accounts. Table 6 shows sample standard questions used by the mock attorney as a framework for each cross examination.

Another ethical consideration can arise in the fact that cross-examination is designed to discredit the witness. Therefore, testifying is a potentially uncomfortable situation. In order to

ensure that discomfort evoked from participants was within reason, the primary investigator, a fifth-year graduate student trained in psychotherapy and risk assessment, monitored each taping session. Clinical judgment determined whether participants approached a degree of discomfort deemed unacceptable. Moreover, a licensed clinical psychologist was available for consultation. If a participant was upset after the testifying experience, identical procedural safeguards were in place to those outlined for moderate levels of self-reported depression. However, no participants reported distress from testifying.

Prior to taping, the mock witness was notified both verbally and in written form (see appendix B) of all the following conditions: a) that approximately six other introductory psychology students, and people they possibly know, will observe these videos, and b) that videos will be kept for a time period of no less than one year in a locked and secured location. Mock witnesses were provided the opportunity to opt out of the study if they were uncomfortable with these conditions. No participants did so. After each session, the participant was debriefed using the debriefing form (see Appendix C) and verbal discussion with the primary investigator and mock attorney.

This phase of the project was pilot tested for the following purposes: a) to ensure that instructions for the mock witnesses were clear, b) to assess whether formulation of cross-examination questions was sufficient, and c) to time the length of each session and decide whether all construct validity questionnaires would be included. For the pilot, each of these aspects was reevaluated after five mock witnesses underwent the entirety of the taping portion of the study. The investigator consulted the dissertation chair regarding these issues; no problems arose during pilot testing.

The final phase was a series of one-hour mock juror sessions; each session consisted of approximately six jurors rating one video. In other words, each testimony video was viewed by a different group of mock jurors. Having the same mock jurors view more than one video may bias their subsequent ratings of witnesses based on a tendency to compare witnesses' testimony and allegation severity.

Each session was guided by a standardized set of procedures. First, participants were notified of their rights as a research participant in both in written and verbal formats. They were also provided the opportunity to ask questions. Mock jurors were then asked if they knew the person testifying while viewing a still video frame. In the two instances this occurred, these participants were asked to wait outside of the room and subsequently dismissed from the session. They still received research credit. Next, participants received questionnaire packets, and subsequently watched a randomly assigned tape of testimony. They then completed the demographics form and ratings of credibility, witness believability, innocence likelihood, juror agreement, and observer WSE. Average scores for each video were tabulated for statistical analyses because the unit of analysis is mock witness, not mock juror. Each mock juror underwent the same consent and debriefing procedures as mock witnesses, only in a group format. As noted earlier, participants were then debriefed in writing and afforded the chance to ask follow-up questions.

RESULTS

Mock Witness Analyses

Cohen, Cohen, West, and Aiken (2003) provide a rationale for choosing statistical analyses. They asserted that simple correlation is best when attempting to clarify the basic relation of two scales, whereas regression is more appropriate to employ when examining a predictive relation (Cohen et al., 2003). Therefore, bivariate correlations were used to assess the WSES with convergent and divergent validity measures completed by mock witnesses. Table 7 displays the correlation matrix of the WSES with convergent and divergent validity measures. Witness age and sex were also included in the table in order to assess potential impact of subject-related variables. These demographic variables produced non-significant findings except that age was correlated with social desirability (r = -.32, p = .05).

Results for mock witness analyses are reported by hypothesis.

H1: Self-reported WSE will demonstrate a non-significant relation with social desirability.

This hypothesis was supported. Witness self-efficacy displayed a non-significant relation with social desirability (r = .26, p = .10).

H2: Self-reported WSE will demonstrate moderate positive relations with social self-efficacy, innocence expectancy, self-confidence, and self-esteem.

This hypothesis was partially confirmed in that two of four hypothesized relations were found. Witness self-efficacy showed moderate positive relations with social self-efficacy (r =

.32, p = .04) and self-confidence (r = .34, p = .03). However, witness self-efficacy displayed non-significant relations with witness innocence expectancy (r = .19, p = .24) and self-esteem (r = .22, p = .16).

H3: Self-reported WSE will demonstrate a strong positive relation with general self-efficacy.

While the direction of this hypothesis was supported, the magnitude was not. Witness self-efficacy showed a moderate positive relation with general self-efficacy (r = .34, p = .03).

H4: Self-reported WSE will demonstrate moderate negative relations with depression and introversion.

This hypothesis was supported for one of two constructs. Witness self-efficacy displayed a moderate negative relation with introversion (r = -.41, p < .01). However, there was a non-significant relation with depression (r = .03, p = .86).

Mock Juror Analyses

Assumptions of parametric statistics were checked prior to all analyses by examining residuals for normality and homoscedasticity. Independence was addressed in overall regression models by examining the aggregate impact of predictor variables in order to ensure that the best predictive model was employed. Control variables of interest were mock witness sex and juror perception of allegation severity. Table 8 is a correlation matrix of control, predictor, and dependent variables for mock juror analyses. Prior to all analyses, predictor variables were standardized.

Results for mock juror analyses are reported by hypothesis.

H5: Self-reported WSE will positively predict mock juror ratings of credibility, agreement, believability, and innocence.

Multivariate regression was used to assess the predictive relations of the WSES with inter-related dependent measures. Average scores on all mock juror ratings were tabulated because the unit of analysis is a mock witness. Thus, there was one overall credibility, innocence likelihood, agreement, believability, and observer WSES rating for each witness. Because these variables were highly inter-correlated (see Table 8), they were entered as dependent measures in the same multivariate regression model. Also, there was one rating of each credibility subscale (confidence, likeability, and trustworthiness) for each witness. The second multivariate model features these inter-related dependent measures. Control variables (witness sex and juror rating of allegation of severity) were not included in these analyses because they were unrelated to all other variables (see Table 8). Effect sizes were measured in partial η^2 values.

This hypothesis was not supported. Witness self-efficacy showed no overall effect on the collection of juror ratings (Wilks' λ [4, 36] = .59, p = .67, partial η^2 = .06). There was no significant impact of WSE on innocence rating (F [1, 39] = 1.36, p = .25), witness believability (F [1, 39] = .19, p = .67), agreement with the witness (F [1, 39] = .15, p = .70), or witness credibility (F [1, 39] = .01, p = .91).

Multivariate analyses assessing witness self-efficacy as a predictor of witness credibility subscales revealed no significant relations (Wilks' λ [4, 36] = .92, p = .36, partial η^2 = .08). There was no significant impact of WSE on witness likeability (F [1, 39] = .48, p = .49), trustworthiness (F [1, 39] = .43, p = .52), or confidence (F [1, 39] = .93, p = .34).

H6: Self-reported WSE will predict variance in mock juror ratings above and beyond confidence and general self-efficacy.

Because WSE was not a significant predictor of dependent measures on its own, it is unlikely that it would show predictive variance above other related constructs such as self-confidence and general self-efficacy. However, WSE was assessed in an additional multivariate regression model in order to test the ability of WSE to predict variance of mock juror ratings in the presence of these covariates. Dependent measures in this model were innocence rating, witness believability, agreement with the witness, and witness credibility. Effect sizes were measured in partial η^2 values.

None of the independent variables were significant predictors of the collection of dependent measures. Self-confidence failed to predict dependent measures (Wilks' λ [4, 34] = .92, p = .46, partial η^2 = .10). There was no significant impact on innocence rating (F [1, 37] = .44, p = .51), witness believability (F [1, 37] = .44, p = .51), agreement with the witness (F [1, 37] = .01, p = .92), or witness credibility (F [1, 37] = .28, p = .60).

General self-efficacy did not predict dependent measures (Wilks' λ [4, 34] = .14, p = .97, partial η^2 = .02). There was no significant impact on innocence rating (F [1, 37] = .01, p = .92), witness believability (F [1, 37] = .05, p = .83), agreement with the witness (F [1, 37] = .02, p = .90), or witness credibility (F [1, 37] = .02, p = .88).

Witness self-efficacy did not predict dependent measures (Wilks' λ [4, 34] = .92, p = .47, partial η^2 = .10). There was no significant impact on innocence rating (F [1, 37] = 1.82, p = .19), witness believability (F [1, 37] = .68, p = .41), agreement with the witness (F [1, 37] = .21, p = .65), or witness credibility (F [1, 37] = .05, p = .82).

The same set of predictor variables was assessed in a multivariate regression with witness credibility subscales as dependent measures. There was no significant effect of independent measures on witness credibility subscales. Self-confidence failed to predict dependent measures

(Wilks' λ [3, 35] = .96, p = .71, partial η^2 = .04). There was no significant impact of self-confidence on witness likeability (F [1, 37] = .99, p = .33), trustworthiness (F [1, 37] = .09, p = .77), or confidence (F [1, 37] = .01, p = .90).

General self-efficacy failed to predict dependent measures (Wilks' λ [3, 35] = 1.00, p = .99, partial η^2 < .001). There was no significant impact of general self-efficacy on witness likeability (F [1, 37] = .02, p = .88), trustworthiness (F [1, 37] = .01, p = .96), or confidence (F [1, 37] = .03, p = .86).

Witness self-efficacy failed to predict dependent measures (Wilks' λ [3, 35] = .94, p = .52, partial η^2 = 06). There was no significant impact of witness self-efficacy on witness likeability (F [1, 37] = .01, p = .93), trustworthiness (F [1, 37] = .18, p = .67), or confidence (F [1, 37] = 1.05, p = .31).

Exploratory Analyses

H7: Self-reported WSE will positively predict observer ratings of WSE.

A simple linear regression was conducted to assess WSE's impact on juror ratings of the witness efficacy. Prior to analyses the same assumptions noted above were assessed; no departures from normality or acceptable levels of equality of variance were found. This hypothesis was not supported. Witness self-efficacy failed to predict juror ratings of efficacy (F [1, 39] = .34, p = .56, R^2 < .01).

The role of observer ratings of WSE in predicting dependent measures

The failure of self-reported WSE to predict dependent measures raises the possibility that accuracy of self-perceptions played a role. Observer (juror) ratings of WSE were originally gathered as a predictive validity criterion for self-reported WSE. The fact that these two constructs are unrelated warrants further attention as well. There is some basis in the self-

efficacy literature related to Axis-I disorders to compare self-reported and observer ratings of self-efficacy (e.g., Doerfler & Aron, 1995; Ducharme & Bachelor, 1993; Guadiano & Herbert, 2007).

Observer ratings of WSE were analyzed in three ways with the goal of clarifying both constructs of self-reported WSE and overall confidence. First, observed WSE was analyzed in a multivariate regression model for its independent predictive ability on innocence rating, agreement with witness, witness believability, and witness credibility. This was followed by a supplemental multivariate regression model predicting witness credibility subscales. Second, observer ratings of WSE were run in the same models with observer ratings of witness confidence (witness credibility subscale) as a covariate. From a theoretical standpoint, this may help differentiate the two constructs where self-reported measures of WSE confidence failed to do so. Third, the interaction between self-reported and observer ratings of WSE was tested in the same multivariate models noted above. These findings may lend insight into whether incongruence between self and observer ratings accounts for null findings for the primary predictive hypotheses.

Multivariate regression analyses showed that observer ratings of witness self-efficacy significantly predicted the collection of dependent measures (Wilks' λ [4, 36] = 6.99, $p < .001$, partial $\eta^2 = .44$). There was a positive trend on innocence rating (F [1, 39] = 3.36, $p = .07$, partial $\eta^2 = .08$). Also, observer ratings of WSE displayed significant positive relations with witness believability (F [1, 39] = 11.19, $p < .01$, partial $\eta^2 = .22$), agreement with the witness (F [1, 39] = 8.65, $p < .01$, partial $\eta^2 = .18$), and witness credibility (F [1, 39] = .23.97, $p < .001$, partial $\eta^2 = .38$).

A follow-up multivariate regression revealed a significant impact of observer ratings of WSE on the collection of witness credibility subscales (Wilks' λ [3, 37] = 13.75, $p < .001$, partial $\eta^2 = .52$). Observer ratings of WSE showed significant positive relations with witness likeability (F [1, 39] = 6.65, $p = .01$, partial $\eta^2 = .15$), trustworthiness (F [1, 39] = 9.68, $p < .01$, partial $\eta^2 = .20$), and confidence (F [1, 39] = 42.03, $p < .001$, partial $\eta^2 = .52$).

Observer ratings of witness confidence versus WSE

Because observer ratings of WSE were more predictive of outcomes than self-reported WSE, it than becomes important to compare observer ratings to theoretically relevant constructs also rated by mock jurors. Of those witness measures originally tested against WSE (i.e., general self-efficacy and self-confidence), only a scale comparable to self-confidence was available from juror ratings (i.e., ratings of witness confidence by jurors). Therefore, observer ratings of witness confidence were analyzed first. Then, observer ratings of WSE and witness confidence were entered together in subsequent multivariate models. Juror ratings of innocence, agreement with the witness, and witness believability were included as dependent measures. Overall witness credibility was left out of the model because witness confidence is included in this total score. Follow-up multivariate regression was conducted to assess the impact observer ratings of WSE and witness confidence on other existing witness credibility subscales, namely witness likeability and trustworthiness.

Observer rated witness confidence emerged as a significant predictor of the collection of dependent measures when entered alone (Wilks' λ [3, 37] = 4.66, $p < .01$, partial $\eta^2 = .27$). Witness confidence showed significant positive associations with innocence likelihood (F [1, 39] = 8.67, $p < .01$, partial $\eta^2 = .18$), agreement with the witness (F [1, 39] = 14.60, $p < .001$, partial $\eta^2 = .27$), and witness believability (F [1, 39] = 12.94, $p = .001$, partial $\eta^2 = .25$).

When both constructs are entered in a multivariate model together, observer ratings of both witness confidence (Wilks' λ [3, 36] = 1.59, p = .21, partial η^2 = .12) and WSE (Wilks' λ [3, 36] = 1.44, p = .25, partial η^2 = .11) show non-significant overall effects. However, observer rated witness confidence displayed significant positive relations with innocence rating (F [1, 38] = 4.84, p = .03, partial η^2 = .11) and agreement with witness (F [1, 38] = 3.94, p = .05, partial η^2 = .09), as well as a positive trend with witness believability (F [1, 38] = 3.69, p = .06, partial η^2 = .09). On the other hand, observer ratings of WSE showed no significant relations innocence rating (F [1, 38] = .06, p = .80, partial η^2 = .01), agreement with the witness (F [1, 38] = 1.03, p = .32, partial η^2 = .03), or witness believability (F [1, 38] = .47, p = .50, partial η^2 = .01). These findings suggest some degree of multi-collinearity or overlap between observed witness confidence and witness self-efficacy.

When observed witness confidence was the only predictor of witness credibility subscales, there was a significant overall effect (Wilks' λ [2, 38] = 8.51, p = .001, partial η^2 = .31). Witness confidence displayed significant positive relations with witness likeability (F [1, 39] = 13.61, p = .001, partial η^2 = .26) and trustworthiness (F [1, 39] = 14.56, p < .001, partial η^2 = .27).

When both observed witness confidence and WSE are entered together predicting witness credibility subscales, a similar pattern emerges as with the collection of overall dependent measures. Observed witness confidence displayed a significant positive trend overall (Wilks' λ [2, 37] = 3.13, p = .06, partial η^2 = .15), whereas observed WSE failed to predict credibility subscales (Wilks' λ [2, 37] = .33, p = .72, partial η^2 = .02). Witness confidence showed significant positive relations with witness likeability (F [1, 38] = 5.82, p = .02, partial η^2 = .13) and trustworthiness (F [1, 38] = 4.14, p = .04, partial η^2 = .10). Although still significant

relations, the magnitude of the associations between witness confidence and other credibility subscales decreases substantially with observed WSE on the model. Observed WSE showed no significant association with witness likeability (F [1, 38] = .03, p = .87, partial η^2 = .001) and trustworthiness (F [1, 38] = .54, p = .47, partial η^2 = .01). Again, these results support the idea that observed witness confidence and WSE are collinear.

Self-reported versus observer ratings of WSE: A moderation analysis

Multivariate regression was used to examine independent effects of self-reported and observed WSE, as well as their interaction, on the collection of primary dependent measures (i.e., innocence rating, agreement with the witness, witness believability, and witness credibility). Self-reported WSE showed no overall effect (Wilks' λ [4, 34] = .68, p = .61, partial η^2 = .07). There was no main effect on innocence rating (F [1, 37] = 2.00, p = .17, partial η^2 = .05), agreement with the witness (F [1, 37] = .50, p = .48, partial η^2 = .01), witness believability (F [1, 37] = .67, p = .42, partial η^2 = .02), or witness credibility (F [1, 37] = .44, p = .51, partial η^2 = .01).

Observed WSE showed a significant main effect on dependent measures (Wilks' λ [4, 34] = 6.89, p < .001, partial η^2 = .45). There was a significant positive trend with innocence rating (F [1, 37] = 3.35, p = .08, partial η^2 = .08), as well as significant positive relations with agreement with the witness (F [1, 37] = 8.02, p < .01, partial η^2 = .18), witness believability (F [1, 37] = 10.52, p < .01, partial η^2 = .22), and witness credibility (F [1, 37] = 23.80, p < .001, partial η^2 = .39).

Although the interaction showed no overall main effect on dependent measures (Wilks' λ [4, 34] = 1.84, p = .15, partial η^2 = .18), a significant interaction emerged for witness credibility (F [1, 37] = 6.04, p = .02, partial η^2 = .14). Figure 1 depicts this interaction. At low levels of

self-reported WSE, all levels of observer WSE are relatively equal in terms of credibility ratings. However, this the gap widens at high levels of self-reported WSE. Witness credibility ratings become markedly lower for individuals with low observed WSE. This linear increase of credibility ratings continues through high levels of observed WSE. The figure appears to show that congruence between high self-reported and high observer ratings of WSE yield high ratings of credibility. Where there is incongruence between self and observer ratings, credibility tends to be lower.

A subsequent multivariate regression model was composed to examine the interaction between self and observer ratings of WSE on credibility subscales. Self-reported WSE showed no overall effect (Wilks' λ [3, 35] = .93, p = .44, partial η^2 = .07). There was no main effect on witness likeability (F [1, 37] = 1.05, p = .37, partial η^2 = .03), trustworthiness (F [1, 37] = 1.16, p = .29, partial η^2 = .03), or confidence (F [1, 37] = .64, p = .43, partial η^2 = .02). Observed WSE showed a significant main effect on witness credibility subscales (Wilks' λ [3, 35] = 14.15, p < .001, partial η^2 = .55). There were significant positive relations trend with witness likeability (F [1, 37] = 6.30, p = .02, partial η^2 = .15), trustworthiness (F [1, 37] = 9.32, p < .01, partial η^2 = .20), and confidence (F [1, 37] = 42.65, p < .001, partial η^2 = .54).

The interaction between self and observer ratings of WSE was significant for witness credibility subscales (Wilks' λ [3, 35] = 2.91, p = .05, partial η^2 = .20). The interaction did not qualify main effects on witness likeability (F [1, 37] = 1.92, p = .18, partial η^2 = .05). However, the interaction qualified main effects on witness confidence (F [1, 37] = 8.04, p < .01, partial η^2 = .18) and trustworthiness (F [1, 37] = 3.16, p = .08, partial η^2 = .08). Figure 2 depicts the moderation of self and observer ratings of WSE on witness confidence. Figure 3 displays the moderation (significant trend) of self and observer ratings of WSE on witness trustworthiness.

Similar patterns to that noted for overall credibility emerge. At low levels of self-reported WSE, all levels of observer WSE are relatively equal in terms of witness confidence and trustworthiness. However, this the gap widens at high levels of self-reported WSE. Witness confidence and trustworthiness ratings become markedly lower for individuals with low observed WSE. This linear increase of credibility subscale ratings continues through high levels of observed WSE. The figure appears to show that congruence between high self-reported and high observer ratings of WSE yield high ratings of witness confidence and trustworthiness. Where there is incongruence between self and observer ratings, witness confidence and trustworthiness tend to be lower.

The moderation between self-reported and observed WSE raise the possibility of a spiked shape relation between WSE and dependent measures. Results from Cramer and colleagues (2009) support this possibility in that they found an inverted U-shaped relation between witness confidence and ratings of credibility. Medium degrees of witness confidence were rated as most credible by mock jurors. For the present study, quadratic fit lines were graphed for self-reported WSE with each of the dependent measures were examined. No inverted U-shaped curves were found.

Inter-rater agreement as an explanation for null findings

Because mock juror hypotheses were largely unsupported, the role of agreement within each jury group was investigated by calculating interclass correlation coefficients (ICCs) for each dependent measure (guilt likelihood, witness believability, witness agreement, witness credibility, and observed WSE). One average ICC was tabulated for each dependent variable. Table 9 contains these results. The ICCs are somewhat low (range .41-.51), even in light of low

sample size. Despite variability one would expect to see within a jury group, these ICCs raise the possibility that inter-rater agreement contributed to null findings.

Summary of Results

Mock witness analyses yielded largely expected convergent and divergent associations. The WSES was significantly positively correlated with general self-efficacy, social self-efficacy, and general self-confidence. The WSES was also significantly negatively related to introversion. The scale displayed non-significant relations with social desirability, depression, self-esteem, and innocence expectancy. The last three relations were contrary to hypotheses.

Mock juror analyses showed little predictive validity for the WSES. When entered as the sole independent variable in multivariate models, the scale failed to predict any of the primary dependent variables, namely innocence rating, agreement with the witness, witness believability, witness credibility, or juror ratings of WSE. Nor did the WSES predict witness credibility subscales (i.e., likeability, confidence, or trustworthiness). These analyses were repeated with general self-efficacy and self-confidence as covariates because they were identified as key theoretical comparative constructs. No significant relations emerged from these analyses.

Exploratory analyses were conducted using observed (juror) ratings of WSE as the primary independent measure. Observer ratings of WSE significantly predicted other primary dependent measures, as well as witness credibility subscales. However, when entered into analyses with observed witness confidence (witness credibility subscale) as a covariate, observed WSE failed to predict any dependent measures. This suggests multi-collinearity between observed WSE and witness confidence.

Further analyses examined the interaction between observed and self-reported WSE ratings. There was a significant interaction between these two variables on witness credibility

ratings, as well as subscales of witness confidence and witness trustworthiness. Figures 1 through 3 depict these results, respectively. Inter-rate agreement within each jury was investigated as a possible explanation for null findings.

Qualitative Observations: The Experience of Testifying

Reflections on the experience of testifying are warranted given the varying reactions participants presented during debriefing. Four thematic categories emerged: witness expectations, witness authenticity, reports regarding components of perceived witness self-efficacy, and witness reactions to cross-examination. Anticipations about the nature of testifying varied based on pre- and post-testimony conversations with witnesses. For instance, many participants indicated having no idea what the experience would be like. Some expected it to be easy, possibly due to the artificial nature of the scenario. However, other witnesses seemed to anticipate a difficult situation that would include being accused or attacked. It would be worth empirically evaluating whether one's expectations about the difficulty or anxiety-provoking nature of testimony influences actual performance.

Witnesses' presence on the stand was quite different as well. Participants' approach to the testifying scenario ranged anecdotally from guarded to extremely open. While most witnesses appeared engaged and enjoyed the experience, a few reacted with indifference or nervousness. Given that most actual witnesses likely report stress and nervousness when testifying, the nature of most participants' reactions may speak to the artificial nature of a psychology 101 project when compared to the real world. The motivation or consequences of partaking in a mock testimony experience versus actual testifying undoubtedly made this project less imposing, and possibly more enjoyable, for participants. These ecological validity concerns not withstanding, witness authenticity, or being oneself on the stand, may be a result of approach

to testifying. Those who were engaged, comfortable, and content likely presented themselves as more authentic than indifferent and nervous counterparts. It appeared that most participants were comfortable and enjoyed the experience of testifying. In light of their comments during debriefing, a few may even have gained some personal insight regarding strengths and weaknesses in strenuous situations.

Much of the informal feedback provided by witnesses confirmed hypothesized facets of WSE. Affective components that witnesses commented on included nervousness and emotional control. Again, most were able to regulate these on the stand in the present study. The important information based on witness feedback is that emotions factor into the experience of testifying. Similarly, witnesses frequently commented on cognitive aspects of testifying such as the ability to recall previous answers or details of the event in the moment of testifying. Based on these observations, testifying may largely be a cognitive experience, especially in relation to organization of one's thoughts on the stand. A final confirming piece of data from witness comments pertains to behavioral components of testifying. Common verbal and non-verbal cues that witnesses conveyed were part of their experience of testifying include, but were not limited to, fidgeting, sitting posture, steadiness of voice, and providing thorough answers. All of these were included in the scale assessing WSE.

Witness reactions to cross-examination were largely consistent. Although the goal of attorney questioning was clearly to ruin witness credibility, most witnesses conveyed a successful ability to cope. In essence, witnesses did not take the cross-examination personally. They understood the adversarial nature, and possibly as a result, were able to frame the experience as something the attorney had to do as part of the experimental task.

DISCUSSION

The overall picture of validity data is mixed. Although convergent and divergent relations for witness self-efficacy were largely supported, expected predictive relations were absent. Confirmation of positive correlations between WSE and other types of self-efficacy and general self-confidence are not surprising. The negative link between introversion and WSE is also logical given prior findings suggesting a negative link between shyness and social self-efficacy (Smith & Betz, 2000). Supported relations with introversion and social self-efficacy support Brodsky's supposition that an element of testifying is akin to story telling or a social situation. Finally, the fact that WSE was unrelated to social desirability across two studies (see also Cramer et al., 2006) is encouraging because mock witnesses may be presenting honest and accurate reflections of their perceived abilities. Given that social desirability was positively related to global constructs (e.g., self-confidence, general self-efficacy), it is also plausible that testifying is an example of a task-specific situation that is intense or influential enough to override one's attempt to present in an overly positive manner.

WSE was unrelated to innocence expectancy, depression, and self-esteem. This contradicts previous general and social self-efficacy literature (e.g., Muris, 2001; Sherer et al., 1981). The simple interpretation of this disparate picture is that WSE is conceptually differentiated from other forms of self-efficacy by expectancy, depression, and self-esteem. A more nuanced look clarifies these differences. While WSE focuses on the ability to testify, innocence expectancy factors in how much the witness either believes their own story or how

others will perceive them. These distinct beliefs may explain the lack of association between outcome expectancy and WSE. The skewed and restricted range of depression obtained in the present sample may have prevented a full examination of the relation with WSE. An equal possibility is that depression may be unrelated to WSE because the former is a fluctuating emotional state rather than a belief system. Similar logic may explain the lack of association with self-esteem. Although self-esteem is a global construct like confidence and general self-efficacy, it centers on self-worth (tied to emotion of depression). Confidence and general self-efficacy largely pertain to assuredness (cognition) and performance (behavior). WSE may be related to global constructs defined by stable thoughts and behavior, as opposed to emotion- and state- dependent constructs.

Regarding predictive validity of WSE, initial findings are clear: Self-reported WSE possesses little value on its own. Not until observer ratings of WSE are factored into predictive models does the construct display any expected predictive validity. Self-reported WSE's failure to predictive outcomes is contrary to self-efficacy data in the organizational psychology literature (e.g., Bauer et al., 2007; Judge et al., 2007; Stajkovic & Luthans, 1998). Two facts lend helpful insights into this discrepancy. First, the organizational psychology literature also shows that complexity of the task can mitigate self-efficacy's ability to predict performance outcome (e.g., Judge et al., 2007; Stajkovic & Luthans, 1998). For example, there is evidence that self-efficacy has more of an effect in lab tasks at medium and high complexity tasks (Stajkovic & Luthans, 1998). The brief nature of the testifying scenario in the present study may not reflect a complex enough task for self-reported WSE to matter. Comparisons of self and other ratings of the same construct also provide information as to why WSE may not matter. Self and observer ratings are not always correlated as is the case in the present study. Congruency between self and observer

ratings is potentially important to consider in assessing the predictive value of WSE. This notion is discussed in further detail below.

Much of the theoretical basis for WSE was based on general and social self-efficacy. Obtaining measures of all three constructs helps to clarify whether these constructs are indeed dissimilar. Overall, WSE was not significantly correlated with either general (r = .34) or social (r = .32) self-efficacy. These are moderate relations, but not statistically significant. Moreover, these three constructs differentially relate to others assessed in mock witnesses. Regarding WSE and general self-efficacy, there were differential relations with social desirability and depression. General self-efficacy was positively correlated with social desirability (r = .46) and negatively related to depression (r = -.35). WSE showed non-significant relations with both of these constructs. General and social self-efficacies also displayed different relations to confidence and self-esteem when compared with WSE. When WSES was compared with the confidence variable, WSES showed a marginal relation (r = .34), while the correlation with general (r = .76) and social (r = .52) self-efficacies were much larger. In short, general and social self-efficacy overlapped more with the construct of confidence. Finally, general (r = .64) and social (r = .40) self-efficacies were significantly related to self-esteem, whereas WSE (r = .22) was not. It appears safe to conclude from these data that WSE is a separate construct from general or social self-efficacy.

Clarifying Self-Efficacy Theory in a Legal Context

The present investigation represents the first examination of Bandura's SET in a psycho-legal context. That said, a theory is validated across situations to the degree that established principles hold in new settings (see Popper, 1959 for a more detailed discussion on theory validation). Based on present results, SET possesses moderate support and requires further

empirical testing in the courtroom. As a starting point, SET appears to show strong support for self-reported construct relations in a legal scenario. WSE was assessed using affective, cognitive and behavioral indicators of the construct. This formulation of self-efficacy is consistent with Bandura's (1997) conceptualization. Present findings that a task-specific belief (WSE) based on Bandura's perspective showed expected relations validates SET. WSE is consistently related to many global traits, which highlights an important premise of SET in a new setting.

Additional issues assessed in the present study offering theoretical clarification include comparisons between a) general self-confidence and self-reported WSE, b) witness confidence and observer ratings of WSE, and c) self and observer ratings of WSE. Each of these predictive evaluations is discussed with attention to theoretical implications.

Bandura (1997) argued that domain-specific constructs are more valuable than global ones for predicting outcomes. The fact that neither general self-confidence nor witness self-efficacy predicted performance outcomes is inconsistent with his argument. When it comes to testifying in court, self-reported perceptions about ability to perform hold little weight. Likewise, what may be considered the closest conceptual global construct to WSE, namely self-confidence, shows limited utility as well. Such is the case for general self-efficacy. Reasons for WSE's lack of validation are numerous. One is that WSE was equally correlated with general self-efficacy and general self-confidence. High conceptual overlap between these constructs may obfuscate predictive analyses. Also worthy of consideration is congruency between self and other ratings of WSE. This is discussed in detail below. Of course, methodological limitations provide possible insight into non-significant findings we well. For instance, the present study had a small and demographically homogenous sample. These limit statistical power and generalizability, respectively. It may be that self-reported WSE or self-confidence would

demonstrate statistically significant relations with a larger sample. Moreover, assessing these constructs in a more representative sample of actual witnesses may produce differing results. Expert witnesses, criminal defendants and police officers are just a few examples of populations that testify more often than undergraduates acting as mock witnesses.

Regarding confidence, a critical literature review (see introduction) raised concern that general self-confidence suffers from lack of a consistent definition and may actually be context-dependent. From a definitional standpoint, Shrauger and Schohn's (1995) general self-confidence used in the present study mirrors general self-efficacy. If their definition of self-confidence is environmentally-dependent, the courtroom does not appear a likely setting conducive to the influence of this construct. High inter-correlations between, general self-confidence, general self-efficacy and WSE also support conceptual overlap between these constructs.

Conclusions regarding SET and self-confidence based on self-report are tempered by additional comparisons involving observer ratings. WSE was compared to another domain-specific belief (witness confidence) based on juror ratings. Although both were independently predictive of performance outcomes, multi-collinearity between observer ratings of these two constructs contradicts tenets of SET. More specifically, confidence is viewed as an ad-hoc judgment from the purview of SET. In regard to testifying, it could be argued that WSE and witness confidence either have much more in common than originally theorized or are in fact the same construct. In this way, SET is not validated with regard to conceptual separation from witness confidence. Indeed, the relation between these constructs remains enigmatic despite the unique setting provided by a courtroom. The role of observer ratings should be viewed as in need of replication given participant limitations of the current study. Although it has been

argued that use of undergraduates is a legitimate way in which to begin testing complex theoretical relations (Bornstein, 1999), observer ratings of WSE may have more of a role in this sample than in more externally valid conditions. Thus, more work using observer ratings of efficacy is needed in order to substantiate how WSE functions in testifying scenarios.

A basic precept of SET is that self-efficacy is an agent of change in and of itself. Various types of self-reported self-efficacy have proven to follow this assertion. Areas where this has held true include academic (e.g., Bandura et al., 2001; Pastorelli et al., 2001), business (e.g., Judge et al., 2007; Stajkovic & Luthans, 1998), and self-defense training (e.g., Ozer & Bandura, 1990) situations to name a few. Witness testimony does not follow suit; therefore, SET does not function as expected in the courtroom setting. The artificial nature of the scenario in the present study may be to blame for this lack of support. Participants' beliefs about testifying may not have produced accurate testimony performance, either due to the lack of a real courtroom setting and personnel, or because of a lack of negative consequences or punishment for poor testimony (e.g., fine, sentencing). Perhaps increasing the verisimilitude between testing conditions of WSE and a real courtroom would produce differing outcomes for the function of SET in this setting.

Another explanation supported by present findings involves the role of observer ratings of WSE. It is not until self-reported WSE is cross-checked against observer perceptions of WSE that the picture is brought into focus. Justification for comparing self and observer ratings exists in the self-efficacy literature (e.g., Doerfler & Aron, 1995; Ducharme & Bachelor, 1993; Guadiano & Herbert, 2007). Observer WSE ratings prove to be a potent predictor of dependent measures apart from self-reported WSE. In the courtroom, SET may only function as an agent of change as it is seen in the eye of the beholder.

Moreover, congruency between self and other perceptions of WSE informs how SET operates in witness testimony, at least in relation to ratings of witness credibility. Where there is congruency between self and observer ratings, there is a positive association with ratings of credibility. This pattern is only true for those who estimated themselves to be high in WSE. This may be so because a third construct explains differences between those with low versus high self-reported WSE. Although it can be argued that confidence may indeed be the explanatory factor, perhaps arrogance, actual ability, or over zealousness offer alternative explanations. For example, it is possible that witnesses displaying arrogance, and who possess actual performance capabilities on the stand, may have been rated highest in credibility. Equally plausible is the notion that witnesses varied in their ability to regulate emotional display on the stand. Those who successfully managed their zeal in testifying may have been perceived accurately by mock jurors, whereas witnesses who failed to regulate emotions about their testimony were likely to have been perceived inaccurately.

From a conceptual standpoint it may prove useful to reassess comparisons between witness self-efficacy and witness self-confidence assumed from the outset. Table 10 contains revisions to assertions originally set forth in Table 3. Bold print reflects conceptual changes based on results from the present study. WSE's basic definition was adjusted to reflect the complexities of its role as an agent of change. Self-reported WSE matters when observer ratings are factored in. Further work is needed to parse out and refine constructs influencing WSE as an agent of change. Likewise, the target of WSE was altered to denote the unequivocal importance of observer ratings.

Regarding witness confidence, two theoretical bases undoubtedly are informative based on present findings. Shrauger and Schohn's (1995) definition of general self-confidence informs

witness theory from the standpoint of self-perceptions. Quite simply, a witness's general self-confidence does not seem to matter for outcomes used in the present study. Witness credibility theory and research now includes a well-formulated conception of witness confidence rated by observers (see Brodsky, Neal, Cramer, & Ziemke, In Press; Cramer et al., 2009; Griffin et al., 2005 for further details). When examined in a predictive testimony scenario, witness confidence defined as a facet of credibility performed well as an independent predictor of other witness outcomes.

Witness self-efficacy and witness confidence showed great statistical overlap. This may be due to the fact that they tap into the same theoretical construct. Therefore, the empirical support for both constructs is framed as mixed or tenuous in order to highlight the need for more research picking apart conditions that may clarify the relation between witness self-efficacy and witness confidence. As a direct result, witness confidence is now presented as a potential target for witness preparation training as well. If the constructs are indeed similar or exactly the same, they can both be examined under the guise of a full training program. This idea is expounded upon below in the discussion of practice applications.

Practical Applications of Witness Self-Efficacy

Just as a theory is validated to the extent its principles hold, a scale is legitimately useful only to the degree it shows associations with relevant constructs. What can be said about the use of WSES? At minimum, the scale can be implemented as a checklist to guide targets for witness preparation because it includes empirically-supported facets of effective testimony. Present findings also suggest that the WSES can be used in self-report and observer forms to assess congruency between perceptions of performance capability on the witness stand. Overall, more

research is needed to strengthen use of the WSES as an outcome measure for witness preparation training. This research is proposed below.

Self-Efficacy in Witness Preparation: An Integrated Approach

In his seminal work on witness preparation research, Boccaccini (2002) advised that new methods for witness preparation are needed. Following this, Boccaccini and colleagues (2004, 2005) proffered the Persuasion Through Witness Preparation (PTWP). As summarized by Neal (2009), The PTWP model centers on bolstering verbal and nonverbal components of testimony including, but not limited to, fidgeting, expressivity, posture, confidence, and emotion. A number of methods inclusive of videotaped feedback are utilized within the scope of PTWP (see Boccaccini 2004, 2005 for more information). The integration of self-efficacy enhancement techniques with the PTWP framework addresses the need for new empirically-based preparation strategies.

The literature on bolstering self-efficacy beliefs is plentiful (e.g., Bandura, 1997; Crain, 2005; Ozer & Bandura, 1990; Schunk & Zimmerman, 2007; Tams, 2008; Yudowitch, Henry, & Gutherie, 2008). As noted earlier, Bandura (1997) outlined four empirically-supported methods for development and adjustment of self-efficacy beliefs. They are actual performance of a behavior, observation of someone performing a behavior, verbal persuasion, and physiological signs.

Each of these concepts either maps on to existing PTWP procedures or offers a complement for use in witness preparation. Consultants and attorneys can draw on all four of these social learning mechanisms of self-evaluation to improve witness self-efficacy and performance. For example, as persons attain more experience on the witness stand, either in actual trials or in mock preparation, they may begin to draw on these actual experiences in order

to develop a strong sense of efficacy on the stand (attempt to perform behavior). The PTWP model already addresses this technique through repeated practice. It may be advantageous to conduct several repetitions with guided feedback whenever possible in order to maximize experience and reduce anxiety.

The education literature also affords some insight into fostering high self-efficacy beliefs and performance capabilities in a format of graduated practice. For example, Schunk and Zimmerman (2007) expound upon self-efficacy literature by outlining a stepwise social-cognitive model of building skills. Their perspective suggests that individuals build skill through the steps outlined by Bandura's approach (e.g., observation, guidance, feedback). The participants then internalize the skill, thereby demonstrating mastery. Only after this process occurs can the skill be utilized in varied circumstances. Applied to witness testimony, PTWP can be used to teach testimony delivery skills. After the witnesses demonstrate mastery under uniform conditions, they can be exposed to changes in order to generalize testimony skills across styles of questioning, settings, and emotional states.

Potential witnesses may also observe an effective model testify successfully and, in turn, incorporate this success by believing they can mimic the behavior (observance/modeling). In essence, the observing witness thinks "If he or she did it, so can I." Bandura (1997) pointed out that modeling is best accomplished when the observer watches someone of similar skill level accomplish a task. Therefore, the witness being trained within PTWP may benefit from watching the prepared testimony of a matched model (i.e., lay person watching another lay person, expert watching another expert). Moreover, use of modeling for witness preparation requires attention to basic tenets of the modeling process: attention, retention, motivation, and reproduction (Bandura, 1997). If trainees do not possess the minimal requisite abilities or motivation to attend

to, retain, and produce effective testimony skills, they will be unable to build high WSE and will perform no better on the stand. Therefore, an integrated PTWP-WSE approach may apply only with witnesses deemed capable of learning via modeling.

A witness's self-efficacy can also be bolstered in the form of reinforcement in witness preparation training, or from positive feedback from mock jurors in trial simulations (verbal persuasion). PTWP modeling can use mock juries or expert raters to provide verbal reinforcement to enhance motivation and belief in one's ability to testify. Likewise, as is suggested in the PTWP model, videotaped feedback can be used to highlight witness successes in learning and applying testifying skills. The self-efficacy based approach to building reading skills offers a compelling example of how reinforcement can be combined with direct observation and modeling (Yudowitch et al., 2008). In short, Yudowitch et al., described a guided approach to building self-efficacy in reading that entailed the following: a) focus on content that can be handled by the trainee, b) establishment of graded, realistic goals, and c) practice of these skills with reinforcement. Extrapolation to witness preparation should focus on a manageable number of skills with graded goals and reinforced practice to maximize witness self-efficacy and performance.

Finally, witnesses can draw conclusions from monitoring physiological cues on the stand such as sweating, steadiness of voice, and muscle tension to self-assess effectiveness their presentation on the stand (that is, through the use of bodily cues). Doing so through practice and training in a variety of settings may help the witness and attorney gauge readiness to testify. For instance, witness training may begin in a private office until physiological cues suggest comfort. Then, the witness may testify in a mock courtroom, or eventually an actual courtroom. Repeated monitoring of physiological cues while on the stand can help assess the state of WSE for the

person being trained. Moreover, simple relaxation techniques such as diaphragmatic breathing can be applied on the stand in order to monitor and control physiological responses, thereby potentially offering another source of high self-efficacy beliefs.

Future work on the development of the WSES may benefit from examination of anxiety (state and trait), overconfidence, and arrogance as convergent and divergent validity constructs. In way of external validity concerns, the WSES can be applied to actual samples to testify, namely criminal defendants, expert witnesses, and police officers. Finally, as outlined in detail above, the WSES can be investigated in a witness preparation training framework to better assess its predictive validity. Indeed, pre-post change scores on the WSES as a result of longitudinal intervention may answer the question of the scale's practical usefulness.

REFERENCES

American Bar Association. (2001). ABA model rules of professional conduct. In *ABA Compendium of Professional Responsibility Rules and Standards* (pp. 1-155). Chicago, IL: American Bar Association.

American Psychological Association. (1992). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author.

Avery, D.R. (2003). Personality as a predictor of the value of voice. *Journal of Psychology: Interdisciplinary and Applied*, *137*, 435-446.

Backenstrass, M., Schwarz, T., Fiedler, P., Joest, K., Reck, C., Mundt, C., & Kronmeuller, K.T. (2006). Negative mood regulation expectancies, self-efficacy beliefs, and locus of control orientation: Moderators or mediators of change in treatment of depression? *Psychotherapy Research*, *16*, 250-258.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191-215.

Bandura, A. (1986). *Social foundations of thoughts and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, *44*, 1175-1184.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *28*, 117-148.

Bandura, A. (1997). *Self-Efficacy: The exercise of control*. New York: W.H. Freeman and Company.

Bandura, A. (2000). Self-Efficacy. In: *Encyclopedia of Psychology, 7*. Kazdin, A.E. Washington: American Psychological Association, 212-213.

Bandura, A. (2005). Guide for constructing self-efficacy scales. In. *Self-Efficacy Beliefs of Adolescents*. Pajares, F., & Urdan, T. (Eds.). Information Age Publishing: Greenich, CT.

Bandura, A., Barbaranelli, C., Caprara, G.V., & Pastorelli, C. (2001). Self-Efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development, 72*, 187-206.

Bauer, T.N., Bodner, T., Erdogan, B., Truxillo, D.M., & Tucker, J.S. (2007). Newcomer adjustment during organizational socialization: A meta-analytic review of antecedents, outcomes, and methods. *Journal of Applied Psychology, 92*, 707-721.

Beck, A.T. (1972). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.

Beck, A.T., Steer, R.A., & Brown, G.K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio: Psychological Corporation.

Beck, A.T., Steer, R.A., Garbin, M.G., (1988) Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*, 77-100.

Beck, A.T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561-571.

Bem, D.J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 74, 183-200.

Bernstein, B.E., & Hartsell, T.L. (2005). *The portable guide to testifying in court for mental health professionals: An a-z guide to being an effective witness*. Hoboken, NJ: John Wiley & Sons, Inc.

Boccaccini, M.T. (2002). What do we really know about witness preparation? *Behavioral Sciences, & the Law*, *20*, 161-189.

Boccaccini, M.T., Brodsky, S.L., & Gordon, T. (2005). Witness preparation training with real and simulated criminal defendants. *Behavioral Sciences & the Law*, *23*, 659-687.

Boccaccini, M.T., Gordon, T, & Brodsky (2004). Effect of witness preparation on witness confidence and nervousness. *Journal of Forensic Psychology Practice*, *3*, 39-51.

Bornstein, B.H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior*, *23*, 75-91.

Braun, K.A., & Loftus, E.F. (1998). Advertising's misinformation effect. *Applied Cognitive Psychology*, *12*, 569-591.

Brenowitz, N., & Tuttle, C.R. (2003). Development and testing of a nutrition-teaching self-efficacy scale for elementary school teachers. *Journal of Nutrition Education Behavior*, *35*, 308-311.

Brewer, W.F., Sampaio, C., & Barlow, M.R. (2005). Confidence and accuracy in the recall of deceptive and nondeceptive sentences. *Journal of Memory and Language*, *52*, 618-627.

Brodsky, S.L. (1991). *Testifying in Court: Guidelines and Maxims for the Expert Witness*. Washington, DC: American Psychological Association.

Brodsky, S.L. (1999). *The Expert Expert Witness: More Maxims and Guidelines for Testifying in Court*. Washington, DC: American Psychological Association.

Brodsky, S.L. (2004). *Coping with Cross Examination and Other Pathways to Effective Testimony*. Washington, DC: American Psychological Association.

Brodsky, S.L., Neal, T.M.S., Cramer, R.J., & Ziemke, M.H. (In Press). Credibility in the courtroom: How likeable should an expert witness be? *Journal of the American Academy of Psychiatry and Law*.

Chen, G., Gully, S.M., & Eden, D. (2001). Validation of a new General Self-Efficacy Scale. *Organizational Research Methods, 4*, 62-83.

Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences (3rd edition)*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Cole, J.G., & McCroskey, J.C. (2003). The association of perceived communication apprehension, shyness, and verbal aggression with perceptions of source credibility and affect in organization and interpersonal contexts. *Communication Quarterly, 51*, 101-111.

Costa, P.T., & McCrae, R.R. (1992). *The NEO-PIR professional manual*. Odessa, FL: Psychological Assessment Resources.

Crain, W.C. (2005). *Theories of Development: Concepts and Applications (5th Edition)*. Upper Saddle Ridge, NJ: Pearson/Prentice Hall.

Cramer, R.J., Brodsky, S.L., & DeCoster, J. (2009). Expert witness confidence and juror extroversion: Their impact on credibility and persuasion in the Courtroom. *Journal of the American Academy of Psychiatry and Law, 37*, 63-74.

Cramer, R.J., Brodsky, S.L., & DeCoster, J. (2006). Initial development of the Witness Self-Efficacy Scale. Unpublished Manuscript. The University of Alabama.

Crowne, D.P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*, 349-354.

Dillon, F.R., & Worthington, R.L. (2003). The Lesbian, Gay, Bisexual Affirmative Counseling Self-Efficacy Inventory (LGB-CSI): Development, validation, and training implications. *Journal of Counseling Psychology*, *50*, 235-251.

Doerfler, L.A., & Aron, J. (1995). Relationship of goal setting, self-efficacy, and self evaluation in dysphoric and socially anxious women. *Cognitive Therapy and Research*, *19*, 725-38.

Ducharme, J., & Bachelor, A. (1993). Perception of social functioning in dysphoria. *Cognitive Therapy and Research*, *17*, 53-70.

Dunn, J.C. (2004). Remember-Know: A matter of confidence. *Psychological Review*, *111*, 524-542.

Dunn, K.F. (2003). Assessing the external validity of jury simulation research: A meta analysis. (Doctoral dissertation, University of Nebraska-Lincoln). *Dissertation Abstracts International*, *63*, 6141.

Everett, S.A., Price, J.A., & Telljohann, S.K. (1996). The elementary-health teaching self-efficacy scale. *American Journal of Health Behavior*, *20*, 90-97.

Gaudiano, B.A., & Herbert, J.D. (2007). Self-efficacy for social situations in adolescents with generalized social anxiety disorder. *Behavioural and Cognitive Psychotherapy*, *35*, 209-223.

Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506-528.

Goldberg, L.R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, &

F. Ostendorf (Eds.), *Personality Psychology in Europe, 7* (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Griffin, M.P., Brodsky, S.L., Blackwood, H., Abboud, B., & Flanagan, C.B. (2005). The Witness Credibility Scale. Unpublished Document. University of Alabama.

Hagler, A.S., Norman, G.J., Zabinski, M.F., Sallis, J.F., Calfas, K.J., & Patrick, K. (2007). Psychosocial correlates of dietary intake among overweight and obese men. *American Journal of Health Behavior, 31,* 3-12.

Haney, C. (1980). Psychology and legal change: On the limits of factual jurisprudence. *Law and Human Behavior, 6,* 191-235.

Haney, C. (1993). Psychology and legal change: The impact of a decade. *Law and Human Behavior, 17,* 371-398.

Healy, M.R., Light, L.L., & Chung, C. (2005). Dual-process models of associativerecognition in younger and older adults: Evidence from receiver characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 768-788.

Hinkle, A.L., Smeltzer, D.J., Allen, C.A., & King, G.D. (1983). The judgments of college students and jurors concerning sanity and guilt of alleged murderers. *The Journal of Social Psychology, 120,* 253-257.

HM Revenue & Customs (2007). AH0880 - Commissioners' hearing: Examination of witnesses: Guidelines for cross examination. http://www.hmrc.gov.uk/manuals/ahmanual/AH0880.htm. Date accessed: February 18, 2007.

International Personality Item Pool (2001). A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences http://ipip.ori.org/. Date accessed: Sept 1, 2006.

Judge, T.A., Jackson, C.L., Shaw, J.C., Scott, B.A., & Rich, B.L. (2007). Self-efficacy and work related performance: The integral role of individual differences. *Journal of Applied Psychology*, *92*, 107-127.

Judge, T.A., Locke, Durham, C.C., & Kluger, A.N. (1998). Dispositional effects on job and life satisfaction: The role of core values. *Journal of Applied Psychology*, *83*, 17-34.

Krauss, D.A., & Sales, B.D. (2001). The effects of clinical and scientific expert testimony on juror decision making in capital sentencing. *Psychology, Public Policy, & Law*, *7*, 267-310.

Loftus, E.F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability in memory. *Learning & Memory*, *12*, 361-366.

Lumpe, A.T., Haney, J.I., & Czerniak, C.M. (2000). Assessing teachers' beliefs about their science teaching context. *Journal of Research in Science Teaching*, *37*, 275-292.

MacCoun, R.J. (1989). Experimental research on jury decision-making. *Science*, *244*, 1046-1050.

McCroskey, J.C. (1966). Scales for the measure of ethos. *Speech Monographs*, *33*, 65-72.

McCroskey, J.C. & Teven J.J. (1999). Good will: A reexamination of the construct and its measurement. *Communication Monographs*, *66*, 90-103.

McCroskey, J.C., & Young, T. (1981). Ethos and credibility: The construct and itsmeasurement after two decades. *Central States Speech Journal*, *32*, 24-34.

Merkle, E.C., & Zandt, T.V. (2006). An application of the Poisson Race Model to confidence calibration. *Journal of Experimental Social Psychology: General*, *135*, 391-408.

Morey, L.C. (1991). *Personality Assessment Inventory - Professional Manual.* Florida: Psychological Assessment Resources, Inc.

Murdock, T.B., Wendler, A.M., & Nilsson, J.E. (2005). Addiction Counseling Self Efficacy Scale (ACSES): Development and initial validation. *Journal of Substance Abuse Treatment*, *29*, 55-64.

Muris, P. (2001). A brief questionnaire for measuring self-efficacy in youths. *Journal of Psychopathology and Behavioral Assessment*, *23*, 145-149.

Neal, T.M.S. (2009). Expert witness preparation: What does the literature tell us? *The Jury Expert*, *21*, 44-52.

O' Barr, W.M. (1982). *Linguistic evidence: Language, power, and strategy in the courtroom*. New York: Academic Press.

Ozer, E.M., & Bandura, A. (1990). Mechanisms governing empowerment effects: A self-efficacy analysis. *Journal of Personality and Social Psychology*, *58*, 472-486.

Pastorelli, C., Caprara, G.V., Barbaranelli, C., Rola, J., Rozsa, S., & Bandura, A. (2001). The structure of children's perceived self-efficacy: A cross-national study. *European Journal of Psychological Assessment*, *17*, 87-97.

Paulhus, D.L. (1991). Measurement and control in response bias; In: *Measures of Personality and Social Psychological Attitudes*. Eds. Robinson, J.P., Shaver, P.R., Wrightsman, L.S.: San Diego, CA: Academic Press, Inc., 1991. 17-60.

Payne, R., & Jahoda, A. (2004). The Glasgow Social Self-Efficacy Scale: A new scale for measuring social self-efficacy in people with intellectual disability. *Child Psychology and Psychotherapy*, *11*, 265-274.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp.123-205). New York: Academic Press.

Popper, K. (1959). *The Logic of Scientific Discovery*. Basic Books: New York City.

Posey, A.J., & Wrightsman, L.S. (2005). *Trial Consulting*. New York, NY: Oxford University Press, 2005.

Pozner, L. (Producer), & Dodd, R. (Producer) (2006). Advanced cross-examination techniques [instructional DVD]. (Available from LexisNexis: http://www.pozneranddodd.com/store/acet_dvd.asp).

Price, P.C., & Stone, E.R. (2004). Intuitive evaluation of likelihood judgment producers: Evidence for a confidence heuristic. *Journal of Behavioral Decision Making*, *17*, 39-57.

Pulford, B., & Colman, A.M. (April, 2005). Testing the confidence heuristic: Are confidence communicators more persuasive? Paper presented at the annual meeting of the British Psychological Society.

Rose, V.G., & Ogloff, J.R.P. (2001). Evaluating the comprehensibility of jury instructions: A method and example. *Law and Human Behavior*, *25*, 409-431.

Rosenberg, M. (1965). *Society and adolescent self-image*. Princeton, NJ: Princeton University Press.

Rosenberg, M. (1986). *Conceiving the Self*. Malabar, FL: Kreiger.

Ryckman, R.M., Robbins, M.A., Thornton, B., & Cantrell, P. (1982). Development and validation of a physical self-efficacy scale. *Journal of Personality and Social Psychology*, *42*, 891-900.

Schunk, D.H., & Zimmerman, B.J. (2007). Influencing children's self-efficacy and self regulation of reading and writing through modeling. *Reading & Writing Quarterly*, *23*, 7-25.

Sechrest, L. (2005). Validity of measures is no simple matter. *Health Services Research*, *40*, 1584-1604.

Severance, L.J., & Loftus, E.F. (1982). Improving the ability of jurors to comprehend and apply criminal jury instructions. *Law & Society Review*, *17*, 153-197.

Severance, L.J., & Loftus, E.F. (1984). Improving criminal justice: Making jury instructions understandable for American jurors. *International Review of Applied Psychology*, *33*, 97-119.

Shaw, J.S., Zerr, T.K., & Woythaler, K.A. (2001). Public eyewitness confidence ratings can differ from those held privately. *Law and Human Behavior*, *25*, 141-154.

Sherer, M., Maddux, J.E., Mercadante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The self-efficacy scale: Construction and validation. *Psychological Reports*, *51*, 663-671.

Sheu, H-B., & Lent, R.W. (2007). Development and initial validation of the Multicultural Counseling Self-Efficacy Scale-Racial Diversity Form. *Psychotherapy: Theory, Research, Practice, Training*, *44*, 30-45.

Shrauger, J.S., & Schohn, M. (1995). Self-confidence in college students: Conceptualization, measurement, and behavioral implications. *Assessment*, *2*, 255-278.

Slovenko, R. (1999). Testifying with Confidence. *Journal of the American Academy of Psychiatry and the Law*, *27*, 127-131.

Smith, H.M., & Betz, N.E. (2000). Development and validation of a scale of perceived social self-efficacy. *Journal of Career Assessment*, *8*, 283-301.

Smith. E.R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108-131.

Soeffing, J.P., Lichstein, K.L., Nau, S.D., McCrae, C.S., Wilson, N.M., Aguillard, R.N., Lester, K.W., & Bush, A.J. (in press). Psychological treatment of insomnia in hypnotic-dependent older adults. *Sleep Medicine.*

Sparks, G.M. (1988). Teachers' attitudes toward change and subsequent improvements in classroom teaching. *Journal of Educational Psychology*, *80*, 111-117.

Sporer, S.L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, & accuracy: A meta-analysis of the confidence-accuracy relationship in eyewitness identification studies. *Psychological Bulletin*, *118*, 315-327.

Stajkovic, A.D., (2006). Development of a core confidence-higher order construct. *Journal of Applied Psychology*, *91*, 1208-1224.

Stajkovic, A.D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, *124*, 240-261.

Steffen, A.M., McKibbin, C., Zeiss, A.M., Gallagher-Thompson, D., & Bandura, A. (2002). The revised scale for caregiving self-efficacy: Reliability and validity studies. *Journal of Gerontology*, *57*, 74-86.

Tams, S. (2008). Constructing self-efficacy at work: A person-centered perspective. *Personnel Review*, *37*, 165-183.

Thomas, J.P., & McFayden, R.G. (1995). The confidence heuristic: A game theoretic analysis. *Journal of Economic Psychology*, *16*, 97-113.

Weitlauf, J.C., Smith R.E., Cervone, D. (2000). Generalization of coping skills training: Influence of self-defense training on women's efficacy beliefs, assertiveness, and aggression. *Journal of Applied Psychology*, *21*, 189-204.

Weitlauf, J.C., Cervone, D., Smith, R.E., Wright, P.M. (2001). Assessing generalization effects in perceived self-efficacy: Multidomain and global assessments of the effects of self defense training for women. *Personality & Social Psychology Bulletin*, *27*, 1683-1691.

Weber, N., & Brewer, N. (2003). The effect of judgment type and confidence scale on confidence-accuracy calibration in face recognition. *Journal of Applied Psychology*, *88*, 490-499.

Weber, N., & Brewer, N. (2004). Confidence-accuracy calibration in absolute and relative face recognition judgments. *Journal of Experimental Psychology: Applied*, *10*, 156-172.

Wells, G.L., Ferguson, T.J., & Lindsay, R.C.L. (1981). The tractability of eyewitness confidence and its implications for triers of fact. *Journal of Applied Psychology*, *66*, 688-696.

Yates, J.F. (1990). *Judgment and decision making*. New York: Prentice Hall.

Yates, J.F., Price, P.C., Lee, J.W., & Ramirez, J. (1996). Good probabilistic forecasters: the 'consumer's' perspective. *International Journal of Forecasting*, *12*, 41-56.

Yudowitch, S., Henry, L.M., & Gutherie, J.T. (2008). Self-efficacy: Building confident readers. In Gutherie, J.T., (Ed.). *Engaging adolescents in reading*. Thousand Oaks, CA: Corwin Press.

Yufang, B. (2004). Compiling the Perceived Academic Self-Efficacy Scale. *Psychological Science*, *27*, 1218-1222.

Table 1

Initial Item Pool for Witness Self-Efficacy Scale (WSES)

Please rate the degree to which you feel you can do the following things if called to testify on the witness stand. Use the scale provided.

1	2	3	4	5
Not well		Moderately well		Very Well

1. Remain calm under cross examination
2. Control my emotions when questioned by an aggressive attorney
3. Be believable
4. Maintain a stable tone of voice when speaking
5. Avoid fidgeting
6. Avoid shifting positions in the chair
7. Acknowledge a degree of certainty in my statements
8. Speak clearly
9. Use vocabulary understandable to members of the jury
10. Respond to questions at a reasonable rate of speech
11. Give explanatory answers to open ended questions
12. Provide descriptive testimony
13. Tell a vivid story to the jury
14. Maintain a good posture throughout the testimony
15. Hold my head up during testimony
16. Avoid slouching
17. Be comfortable on the witness stand
18. Remain poised when being questioned by an attorney
19. Maintain eye contact with the jury
20. Hold eye contact with an attorney
21. Pay attention to questions asked by the attorney
22. Remember all the details of my previous answers
23. Recall facts of the case
24. Ask for a question to be repeated if I do not remember what was said
25. Hide my nervousness
26. Convey confidence in my ability
27. Not get lost in my own thoughts when testifying
28. Organize my thoughts
29. Speak up to the judge when I feel my rights have been violated
30. Use appropriate inflections in my voice when answering questions
31. Keep myself from making unnecessary gestures
32. Limit my bodily movements on the stand
33. Comfortably admit when I am uncertain of an answer
34. Speak so all members of the court can understand my responses
35. Sit up

36. Lean slightly forward when answering some questions
37. Provide more than “yes/no” answers
38. Persuade members of the jury
39. Appear credible
40. Act natural
41. Be myself when testifying
42. Remain relaxed

Table 2

Initial Construct Validity for WSES

	Social Desirability	*General Self-Efficacy*	*Social Self-Efficacy*	*Witness Self-Efficacy*
Social Desirability	-	-.38**	-.25**	.03
General Self-Efficacy		-	.61**	.40**
Social Self-Efficacy			-	.30**
Witness Self-Efficacy				-

Note: ** $p < .01$

Table 3

Comparisons of Self-Efficacy, Confidence, Witness Self-Efficacy, and Witness Confidence

Factor	Self-Efficacy	Confidence	Witness Self-Efficacy	Witness Confidence
Definition	Affirmation of ability *and* strength of belief	*Only* degree of certainty in outcome	Belief in actual ability to testify and agent of change for testimony	Belief in degree of certainty in responses on the stand
Components	Behavioral, cognitive, and affective	Cognitive and affective (Inconsistent on behavioral)	Behavioral, cognitive, and affective	Cognitive and affective
Target	Specific behaviors prior to action	Judgments resulting from action	Empirically-supported efficacious behaviors	Self or other judgment of effectiveness on the witness stand
Theoretical Basis	Social-Cognitive and Self-Efficacy Theories	Fragmented	General, social, and teaching self-efficacy principles applied to the law	None
Empirical Support	Considerable amount across areas	Considerable amount across areas	One unpublished study	Many studies based on raters judgments of witness confidence
Utility	Belief system acting as agent of change; can be a target of intervention	Construct that results from intervention	Potential measure and target for witness preparation training	Measure for witness credibility research

Table 4

Descriptive Statistics and Psychometric Properties of Measures by Section of Study

Sample/Measure	*Mean*	*SD*	*Range*	*α*
Mock Witness Measures (N = 41)				
Witness Self-Efficacy	115.80	16.94	81.00-146.00	.94
Social Desirability	16.46	5.46	5.00-26.00	.94
Social Self-Efficacy	22.32	3.83	13.00-29.00	.88
General Self-Efficacy	65.90	10.26	44.00-85.00	.80
Self-Confidence	20.10	3.37	13.00-28.00	.65
Self-Esteem	33.07	4.61	22.00-40.00	.74
Introversion	20.70	6.88	10.00-36.00	.82
Depression	6.17	5.11	0.00-21.00	.86
Witness Innocence Expectancy	7.75	1.74	3.00-10.00	N/A
Mock Juror Measures (N = 289)				
Witness Credibility	98.36	21.60	37.00-148.00	.94
Witness Confidence	32.71	8.51	7.00-50.00	.92
Witness Trustworthiness	28.65	9.01	5.00-50.00	.96
Witness Likeability	37.00	8.25	5.00-50.00	.89
Allegation Severity	3.64	2.00	1.00-10.00	N/A
Guilt Likelihood	5.30	2.52	1.00-10.00	N/A
Agreement with Witness	5.25	2.35	1.00-10.00	N/A
Witness Believability	5.30	2.38	1.00-10.00	N/A
Observed Witness Self-Efficacy	105.64	21.78	53.00-149.00	.95

Table 5

Type of Self-Reported Offenses for Mock Witnesses

Type of Allegation/Offense	Frequency
Damaging property (e.g., car accident)	10
Academic Misconduct (e.g., plagiarism)	6
Relationship infidelity	6
School misbehavior (e.g., talking in class)	5
Other misbehavior (e.g., breaking curfew)	5
Gossiping	4
Automotive citations	3
Lying	2

Table 6

Sample Questions Used by Mock Attorney

"Would you please state your name for the record?"

"Mr./Ms. X, my name is Tess Neal and I am the cross-examining attorney for this case. It is my understanding that you were accused of (insert accusation). Do you agree this is what you were accused of?"

"Do you agree to testify openly and honestly about these accusations?"

"You understand that you are under cross-examination and are expected to tell the truth, the whole truth, and nothing but the truth?"

"Tell me about what happened that led to the accusation."

"Did you do (insert accusation)?"
Follow-up question: "Have you ever lied before, Mr./Ms. X?",

Sample accusation-relevant questions:
“Isn’t it true you have consumed alcohol under before?”
“Isn’t it true you got in trouble with this teacher before?
“Isn’t it true you admitted to cheating on your girlfriend before?”

"Let me make sure I've got all this right. (Insert a summary of the allegation and testimony, with emphasis on particularly damaging admissions [e.g., you admitted to this court that you've lied before, and now you want this court to believe you when you say you didn't do [insert allegation]). Is that correct?"

"Thank you, Mr./Ms. X, I have no further questions for you."

Table 7

Correlation Matrix of WSES, Convergent Validity, and Divergent Validity Measures

	1	2	3	4	5	6	7	8	9	10	11
1. Age	-	-.19	.16	.04	-.32*	.12	-.06	.28+	.14	-.02	-.10
2. Sex		-	.10	-.18	.27	.11	.11	.07	-.09	.15	-.02
3. WIE			-	.19	-.02	-.15	-.06	-.24	-.28+	.25	-.03
4. WSES				-	.26	.34*	.32*	.34*	.22	-.03	-.41**
5. SDS					-	.46**	.09	.27+	.39*	-.18	.17
6. GSES						-	.55***	.76***	.64***	-.35*	-.40**
7. SES							-	.52***	.40**	-.24	-.77***
8. PEI-G								-	.70***	-.35*	-.40**
9. RSES									-	-.56***	-.29+
10. BDI										-	.20
11. Intro											-

*Note: Correlations reported in r values; + = p < .10; * = p < .05; ** p < .01; *** p < .001.*
WIE = Witness Innocence Expectancy; WSES = Witness Self-Efficacy; SDS = Social Desirability; GSES = General Self-Efficacy; SES = Social Self-Efficacy; PEI-G = General Self-Confidence; RSES = Self-Esteem; BDI = Depression; Intro = Introversion.

Table 8

Correlation Matrix of WSES and Mock Juror Measures

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Sex	-	.02	-.18	.11	.07	.15	.06	.10	.10	.18	.26	-.18	-.18
2. AS		-	.21	.14	.22	.02	.09	.12	.16	.07	.19	.14	.11
3. WSES			-	.34*	.34*	-.18	.07	.06	.02	.11	.10	.15	.09
4. GSES				-	.76***	.03	.13	.03	-.12	-.23	-.09	-.01	.18
5. PEI-G					-	.07	.16	.02	-.15	-.28+	-.11	-.01	.08
6. Verdict						-	.84***	.82***	.49***	.30+	.52***	.43**	.28+
7. Believe							-	.93***	.66***	.42**	.73***	.52***	.47**
8. Agree								-	.71***	.53***	.79***	.50***	.43**
9. Cred									-	.86***	.87***	.81***	.62***
10. Like										-	.72**	.51**	.38*
11. Trust											-	.52***	.45**
12. Conf												-	.72***
13.ObWSE													-

Note: Correlations reported in r values; + = p < *.10*; * = p < *.05*; ** p < *.01*; *** p < *.001*.
As = Allegation Severity; WSES = Witness Self-Efficacy; GSES = General Self-Efficacy; PEI-G = Self-Confidence; Verdict = Likelihood of Finding the Witness Not Guilty; Believe = Witness Believability; Agree = Agreement with Witness; Cred = Witness Credibility; Like = Witness Likeability; Trust = Witness Trustworthiness; Conf = Witness Confidence; ObWSE = Juror Rating of Witness Efficacy.

Table 9

Average Inter-rater Agreement for Mock Jury Groups by Dependent Measure

Dependent Variable	*Average ICC value*
Guilt Likelihood	.51
Witness Believability	.45
Witness Agreement	.41
Witness Credibility	.50
Observed WSE	.44

Table 10

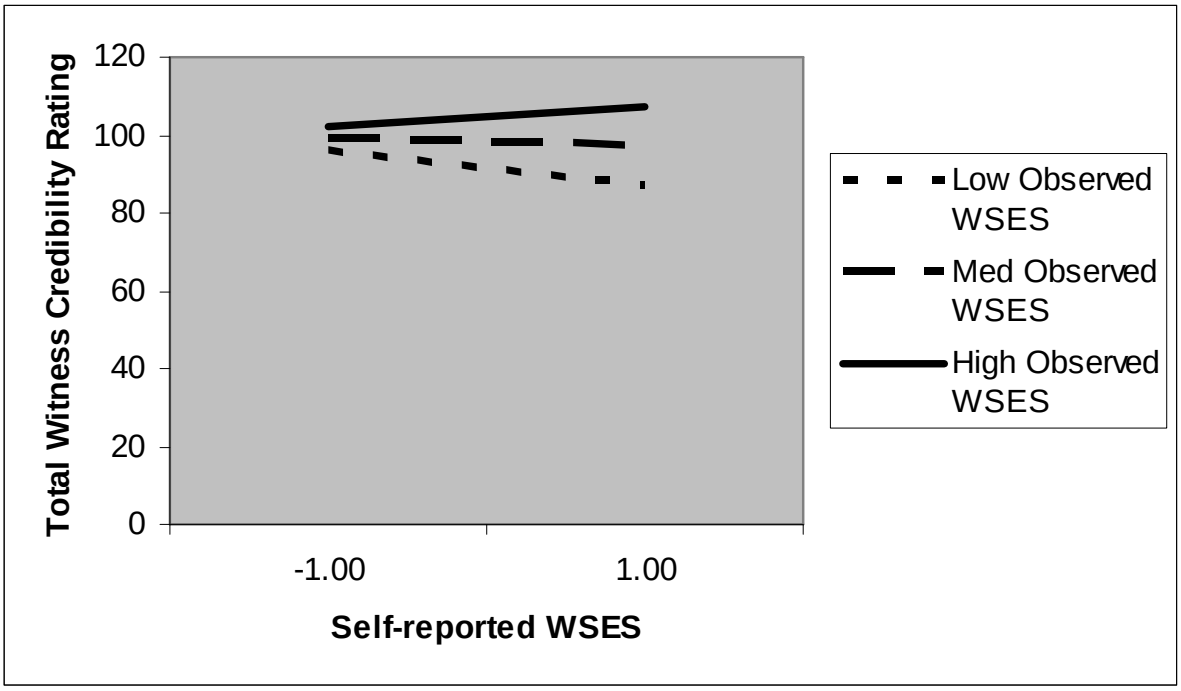
Revised Comparison of Witness Self-Efficacy and Witness Confidence

Factor	Witness Self-Efficacy	Witness Confidence
Definition	Belief in actual ability to testify, **but currently lacking support as an agent of change for testimony**	Belief in degree of certainty in responses on the stand
Components	Behavioral, cognitive, and affective	Cognitive and affective
Target	**Self or other judgments about** empirically-supported efficacious behaviors	Self or other judgment of effectiveness on the witness stand
Theoretical Basis	General, social, and teaching self-efficacy principles applied to the law	**Witness Credibility Theory; General Self-Confidence Theory**
Empirical Support	One unpublished study; **Mixed results from the present study**	Many studies based on raters judgments of witness confidence; **Mixed results from the present study**
Utility	Potential measure and target for witness preparation training	Measure for witness credibility research, **and potential for witness preparation training**

Note: **Bold print** indicates changes to the Table based on results of the present study.

Figure 1

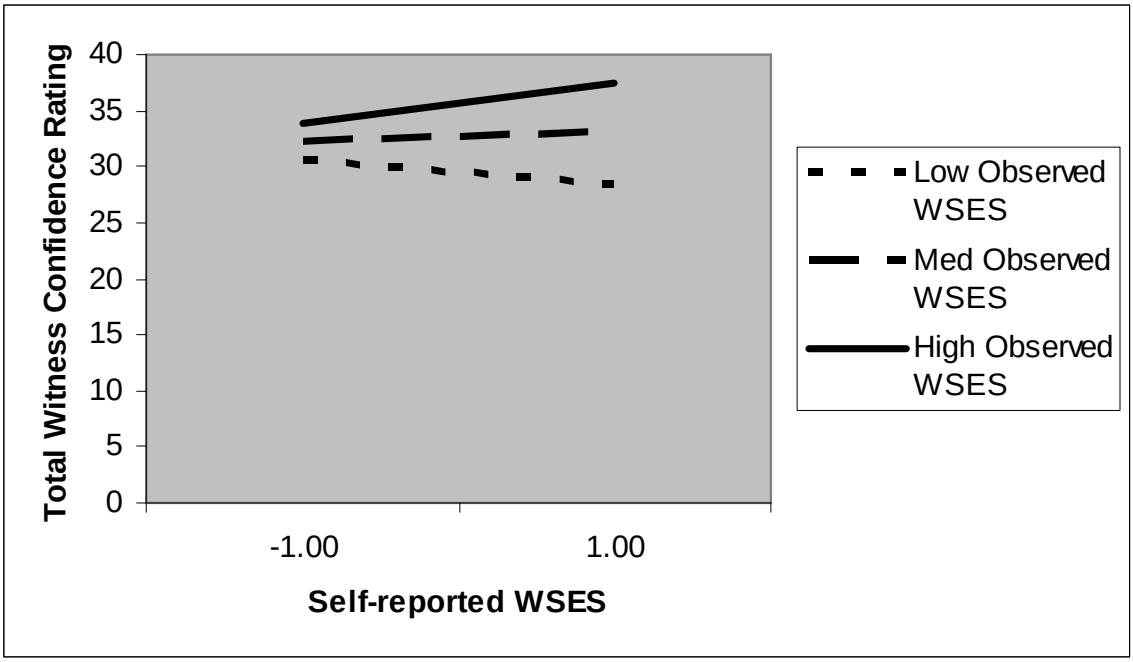
Self-Reported WSES by Observed WSES on Overall Witness Credibility



Note: -1 = Low Self-reported WSES; 1 = High Self-reported WSES

Figure 2

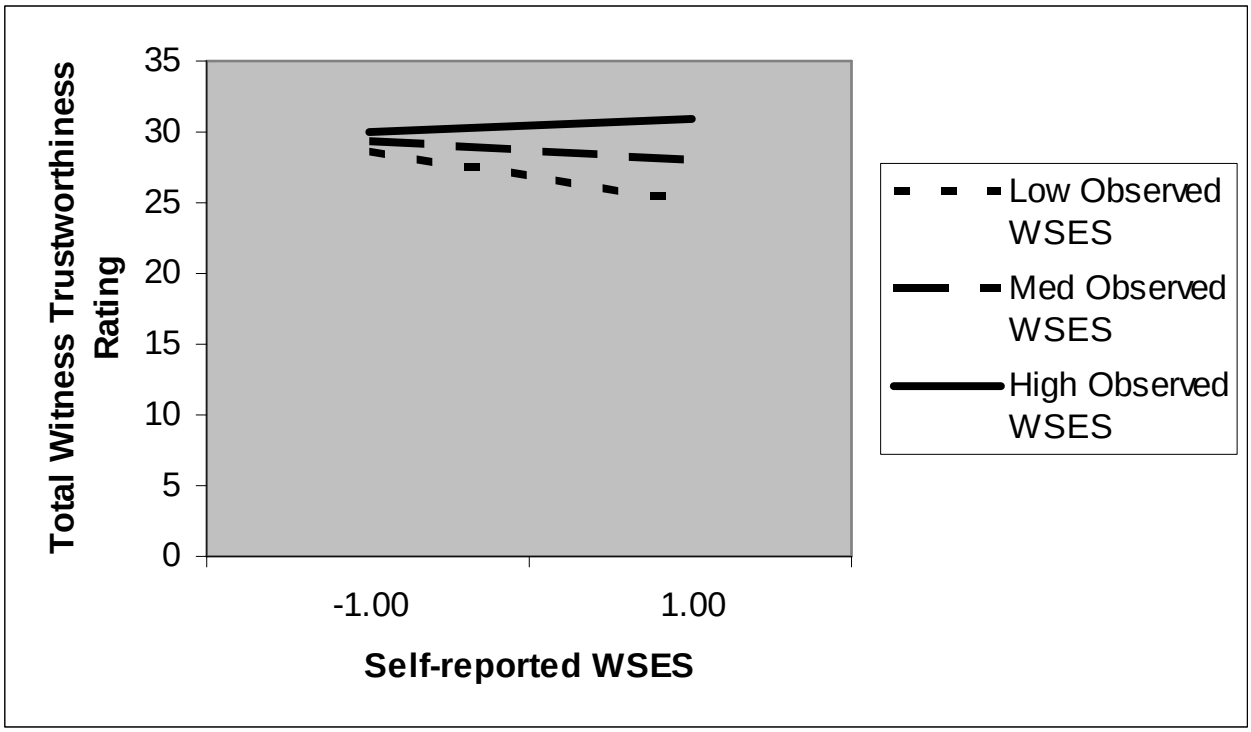
Self-Reported WSES by Observed WSES on Overall Witness Confidence Ratings



Note: -1 = Low Self-reported WSES; 1 = High Self-reported WSES

Figure 3

Self-Reported WSES by Observed WSES on Overall Witness Trustworthiness Ratings (Trend)



Note: -1 = Low Self-reported WSES; 1 = High Self-reported WSES

APPENDIX A: MEASURES

Please provide the following information:

Age: _____

Sex: M F

Ethnicity (Check all that apply):
Caucasian _____
African-American _____
Latin-American _____
Asian-American _____
Other (please specify) _____

Please identify your religion (if any): _____

What is your mother's (or other guardian) occupation: _____

What is your father's (or other guardian) occupation: _____

Have you ever testified in court? YES NO

If yes, for what reason? _____
_____

Have you ever been called for jury duty? YES NO

If yes, have you ever served on a jury in a trial? YES NO

If yes, for what type of trial and what was the verdict? Please briefly explain: _____
_____

WSES

Please rate the degree to which you feel **you** can do the following things if called to testify on the witness stand. Use the scale provided.

1	2	3	4	5
Not well		Moderately well		Very Well

1. Remain calm under cross examination _____
2. Control my emotions when questioned by an aggressive attorney _____
3. Maintain a stable tone of voice when speaking _____
4. Avoid fidgeting _____
5. Speak clearly _____
6. Use vocabulary understandable to members of the jury _____
7. Respond to questions at a reasonable rate of speech _____
8. Give explanatory answers to open ended questions _____
9. Tell a vivid story to the jury _____
10. Maintain a good posture throughout the testimony _____
11. Be comfortable on the witness stand _____
12. Remain poised when being questioned by an attorney _____
13. Maintain eye contact with the jury _____
14. Hold eye contact with an attorney _____
15. Pay attention to questions asked by the attorney _____
16. Recall facts of the case _____
17. Ask for a question to be repeated if I do not remember what was said _____
18. Hide my nervousness _____
19. Convey confidence in my ability _____
20. Organize my thoughts _____
21. Speak up to the judge when I feel my rights have been violated _____
22. Use appropriate inflections in my voice when answering questions _____
23. Comfortably admit when I am uncertain of an answer _____
24. Speak so all members of the court can understand my responses _____
25. Sit up _____
26. Lean slightly forward when answering some questions _____
27. Provide more than “yes/no” answers _____
28. Appear credible _____
29. Act natural _____
30. Be myself when testifying _____

MCSDS

Listed below are a number of statements concerning personal attitudes and traits. Read each item and decide whether the statement is true or false as it pertains to you.

T F 1. Before voting I thoroughly investigate the qualifications of all the candidates.

T F 2. I never hesitate to go out of my way to help someone in trouble.

T F 3. It is sometimes hard for me to go on with my work if I am not encouraged.

T F 4. I have never intensely disliked anyone.

T F 5. On occasion I have had doubts about my ability to succeed in life.

T F 6. I sometimes feel resentful when I don't get my way.

T F 7. I am always careful about my manner of dress.

T F 8. My table manners at home are as good as when I eat out in a restaurant.

T F 9. If I could get into a movie without paying and be sure I was not seen, I would probably do it.

T F 10. On a few occasions, I have given up doing something because I thought too little of my ability.

T F 11. I like to gossip at times.

T F 12. There have been times when I felt like rebelling against people in authority even thought I knew they were right.

T F 13. No matter who I'm talking to, I'm always a good listener.

T F 14. I can remember "playing sick" to get out of something.

T F 15. There have been occasions when I took advantage of someone.

T F 16. I'm always willing to admit it when I make a mistake.

T F 17. I always try to practice what I preach.

T F 18. I don't find it particularly difficult to get along with loud-mouthed, obnoxious people.

T F 19. I sometimes try to get even, rather than forgive and forget.

T F 20. When I don't know something I don't at all mind admitting it.

T F 21. I am always courteous, even to people who are disagreeable.

T F 22. At times I have really insisted on having things my own way.

T F 23. There have been occasions when I felt like smashing things.

T F 24. I would never think of letting someone else be punished for my wrongdoings.

T F 25. I never resent being asked to return a favor.

T F 26. I have never been irked when people expressed ideas very different from my own.

T F 27. I never make a long trip without checking the safety of my car.

T F 28. There have been times when I was quite jealous of the good fortune of others.

T F 29. I have almost never felt the urge to tell someone off.

T F 30. I am sometimes irritated by people who ask favors of me.

T F 31. I have never felt that I was punished without cause.

T F 32. I sometimes think when people have a misfortune they only got what they deserved.

T F 33. I have never deliberately said something that hurt someone's feelings.

SES

Instructions. This questionnaire is a series of statements about your personal attitudes and traits. Each statement represents a commonly held belief. Read each statement and decide to what extent it describes you. There are no right or wrong answers. You will probably agree with some of the statements and disagree with others, Please indicate your own personal beliefs about each statement below by marking the letter that best describes your attitude or feeling. Please be truthful and describe yourself as you really are, not as you would like to be.

Mark:

A If you **DISAGREE STRONGLY** with the statement
B If you **DISAGREE MODERATELY** with the statement
C If you **neither agree or disagree** with the statement
D If you **AGREE MODERATELY** with the statement
E If you **AGREE STRONGLY** with the statement

1. I like to grow houseplants. _____
2. When I make plans, I am certain I can make them work. _____
3. One of my problems is that I cannot get down to work when I should. _____
4. If I can't do the job the first time, I keep trying until I can. _____
5. Heredity plays the major role in determining one's personality. _____
6. It is difficult for me to make new friends. _____
7. When I set important goals for myself, I rarely achieve them. _____
8. I give up on things before completing them. _____
9. I like to cook. _____
10. If I see someone I would like to meet, I go to that person instead of waiting for him or her to come to me. _____
11. I avoid facing difficulties. _____
12. If something looks too complicated, I will not even bother to try it. _____
13. There is some good in everybody. _____
14. If I meet someone interesting who is very hard to make friends with, I'll soon stop trying to make friends with that person. _____
15. When I have something unpleasant to do, I stick to it until I finish it. _____

16. When I decide to do something, I go right to work on it. _____

17. I like science. _____

18. When trying to learn something new, I soon give up if I am not initially successful. _____

19. When I'm trying to become friends with someone who seems uninterested at first, I don't give up very easily. _____

20. When unexpected problems occur, I don't handle them very well. _____
21. If I were an artist, I would like to draw children. _____

22. I avoid trying to learn new things if they look too difficult for me. _____

23. Failure just makes me try harder. _____

24. I do not handle myself well in social gatherings. _____

25. I very much like to ride horses. _____

26. I feel insecure about my ability to do things. _____

27. I am a self-reliant person. _____

28. I have acquired my friends through my personal abilities at making friends. _____

29. I give up easily. _____

30. I do not seem capable of dealing with most problem

31. I can handle situations that come up in my life. _____

Witness Innocence Rating

***Instructions**: Answer the following question on a 10-point scale. Please circle your answer.*

Based upon your testimony, how likely do you think the jury is to find you **innocent**?

10	9	8	7	6	5	4	3	2	1
Extremely likely			Very likely		Likely		Not likely	Very unlikely	

PEI-G

Below are listed a number of statements that reflect common feelings, attitudes, and behaviors. Please read each statement carefully and think about whether you agree or disagree that it applies to you. Try to respond *honestly* and *accurately*, but it is not necessary to spend much time deliberating about each item. Think about how the item applies to you during the last two months unless some other time period is specified. Indicate your degree of agreement with each statement in the space next to the number using the following scale:

A	B	C	D
STRONGLY AGREE	MAINLY AGREE	MAINLY DISAGREE	STRONGLY DISAGREE

1. I often feel unsure of myself even in situations I have successfully dealt with in the past. _____
2. I lack some important capabilities that may keep me from being successful. _____
3. Much of the time I don't feel competent as many of the people around me. _____
4. I have fewer doubts about my abilities today than most people. _____
5. When things are going poorly, I am usually confident that I can successfully deal with them. _____
6. I have more confidence in myself than most people I know . _____
7. If I were more confident about myself, my life would be better. _____

RSES

Instructions: Below is a list of statements dealing with your general feelings about yourself. If you strongly disagree, circle **SA**. If you agree with the statement, circle **A**. If you disagree, circle **D**. If you strongly disagree, circle **SD**.

1. On the whole, I am satisfied with myself.	**SA**	**A**	**D**	**SD**
2. At times, I think I am no good at all.	**SA**	**A**	**D**	**SD**
3. I feel that I have a number of good qualities.	**SA**	**A**	**D**	**SD**
4. I am able to do things as well as most other people.	**SA**	**A**	**D**	**SD**
5. I feel I do not have much to be proud of.	**SA**	**A**	**D**	**SD**
6. I certainly feel useless at times.	**SA**	**A**	**D**	**SD**
7. I feel I am a person of worth, at least on an equal plane with others.	**SA**	**A**	**D**	**SD**
8. I wish I could have more respect for myself.	**SA**	**A**	**D**	**SD**
9. All in all, I am inclined to feel I am a failure.	**SA**	**A**	**D**	**SD**
10. I take a positive attitude toward myself.	**SA**	**A**	**D**	**S**

IPIP

Instructions: Please indicate how true each statement below is of you on a regular basis using the following 5-point scale:
1 = “Not at all like me”
2 = “Somewhat like me”
3 = “Like me”
4 = “Very much like me”
5 = “Extremely like me”
_____ 1. Feel comfortable around people
_____ 2. Have little to say
_____ 3. Make friends easily
_____ 4. Keep in the background
_____ 5. Am skilled in handling social situations
_____ 6. Would describe my experiences as somewhat dull
_____ 7. Am the life of the party
_____ 8. Don’t like to draw attention to myself
_____ 9. Know how to captivate people
_____ 10. Don’t talk a lot

WCS

Instructions: *Please rate the* ***witness*** *for the following items on the scale provided.*
If you are unsure, please take your BEST GUESS.

Example:

1	*2*	*3*	*4* ✓	*5*	*6*	*7*	*8*	*9*	*10*
Dressed Formally									*Dressed Informally*

1	2	3	4	5	6	7	8	9	10

Unfriendly Friendly

1	2	3	4	5	6	7	8	9	10

Disrespectful Respectful

1	2	3	4	5	6	7	8	9	10

Unkind Kind

1	2	3	4	5	6	7	8	9	10

Ill-mannered Well-mannered

1	2	3	4	5	6	7	8	9	10

Unpleasant Pleasant

1	2	3	4	5	6	7	8	9	10

Untrustworthy Trustworthy

1	2	3	4	5	6	7	8	9	10

Untruthful Truthful

1	2	3	4	5	6	7	8	9	10

Undependable Dependable

1	2	3	4	5	6	7	8	9	10

Dishonest Honest

1	2	3	4	5	6	7	8	9	10

Unreliable Reliable

1	2	3	4	5	6	7	8	9	10

Not confident — Confident

1	2	3	4	5	6	7	8	9	10

Inarticulate — Well-spoken

1	2	3	4	5	6	7	8	9	10

Tense — Relaxed

1	2	3	4	5	6	7	8	9	10

Shaken — Poised

1	2	3	4	5	6	7	8	9	10

Not Self-Assured — Self-Assured

1	2	3	4	5	6	7	8	9	10

Uninformed — Informed

1	2	3	4	5	6	7	8	9	10

Illogical — Logical

1	2	3	4	5	6	7	8	9	10

Uneducated — Educated

1	2	3	4	5	6	7	8	9	10

Unwise — Wise

1	2	3	4	5	6	7	8	9	10

Unscientific — Scientific

Juror Rating of Allegation Severity

***Instructions**: Answer the following question on a 10-point scale. Please circle your answer.*

How severe do you consider the allegation against the defendant?

10	9	8	7	6	5	4	3	2	1
Extremely severe		Very Severe		Moderately Severe		Mostly harmless		Harmless	

Juror Rating of Guilt and Agreement

***Instructions**: Answer the following question on a 10-point scale. Please circle your answer.*

Based upon the witness's testimony:

How likely are you to find the defendant **not guilty**?

10	9	8	7	6	5	4	3	2	1
Extremely likely		Very likely		Likely		Not likely		Very unlikely	

How likely are you to **believe the defendant's testimony**?

10	9	8	7	6	5	4	3	2	1
Extremely likely		Very likely		Likely		Not likely		Very unlikely	

How likely are you to **agree with the defendant's story**?

10	9	8	7	6	5	4	3	2	1
Extremely likely		Very likely		Likely		Not likely		Very unlikely	

Juror WSES

Please rate the degree to which you feel **the witness** did do the following things if called to testify on the witness stand. Use the scale provided.

1	2	3	4	5
Not well		Moderately well		Very Well

1. Remain calm under cross examination _____
2. Control their emotions when questioned by an aggressive attorney _____
3. Maintain a stable tone of voice when speaking _____
4. Avoid fidgeting _____
5. Speak clearly _____
6. Use vocabulary understandable to members of the jury _____
7. Respond to questions at a reasonable rate of speech _____
8. Give explanatory answers to open ended questions _____
9. Tell a vivid story to the jury _____
10. Maintain a good posture throughout the testimony _____
11. Be comfortable on the witness stand _____
12. Remain poised when being questioned by an attorney _____
13. Maintain eye contact with the jury _____
14. Hold eye contact with an attorney _____
15. Pay attention to questions asked by the attorney _____
16. Recall facts of the case _____
17. Ask for a question to be repeated if I do not remember what was said _____
18. Hide nervousness _____
19. Convey confidence in their ability _____
20. Organize their thoughts _____
21. Speak up to the judge when they feel their rights have been violated _____
22. Use appropriate inflections in their voice when answering questions _____
23. Comfortably admit when they are uncertain of an answer _____
24. Speak so all members of the court can understand my responses _____
25. Sit up _____
26. Lean slightly forward when answering some questions _____
27. Provide more than “yes/no” answers _____
28. Appear credible _____
29. Act natural _____
30. Be themselves when testifying _____

APPENDIX B: MOCK DEFENDANT NOTIFICATION FORM

To the participant:

You have elected to participate in a study pertaining to witness testimony. Conditions of this study require that:
a) Six other py101 students will observe these videos for research purposes. These may be people who know you.
b) Videos will be kept by the researcher for a time period of no less than one year in a locked and secured location.

If you are uncomfortable with either of these conditions you may withdraw from this study without penalty or judgment. Indeed, it is the goal of the primary investigator to address possible concerns and avoid harm to the participant in accordance with American Psychological Association ethical guidelines and the University of Alabama Human Subjects Review Board. If you have any questions or concerns, you can reach the research supervisor, Dr. Stanley Brodsky at (205) 348-1920, or the primary investigator, Robert J. Cramer via email at Crame001@bama.ua.edu. If you agree to these conditions, please sign and date below:

_____ _____
Participant Signature Date

_____ _____
Researcher Signature Date

Thank you for your participation,

Robert J. Cramer, M.A.

APPENDIX C: PARTICIPANT INFORMATION SHEET AND DEBRIEFING FORM

Research Study Information Sheet
Title of Research Project: Testifying in Court

Investigators: Robert J. Cramer, M.A., Stanley L. Brodsky, PhD
It is important that you read the following explanation of this research study. This document describes the purpose, procedures, possible benefits and risks, and confidentiality of this study.

Purpose and Procedures
The current study is examining the testimony of mock witnesses and decision-making process of mock-jurors. If you decide to be in this study and serve as a witness, you will be asked to provide written accounts of past accusations against you, complete a series of questionnaires, and undergo videotaped testimony. This will take roughly two hours. If you decide to serve as a juror in this study, you will watch a video of testimony, rate the expert on adjectives, and complete a few questionnaires by paper and pencil. Being in this study will take about 45 minutes.

Benefits and Risks
There are no direct benefits to you for participating in the study, but you will receive one (1) or two (2) research credits depending on which stage of the study you participate in. Potential benefits to you are gaining insight into your personal beliefs regarding justice and receiving information about the judicial process. This study will help psychologists and lawyers better understand the decision-making processes of mock-jurors as it relates to witness testimony. There are no foreseeable risks or discomforts involved with participating in this study other than those detailed in the defendant notification form.

Confidentiality
Your name will only be recorded to ensure you receive credit in your course for your participation and will be kept separate from the other study materials. The documents containing the names of participants will be destroyed once all credit has been given for participation. There will be no identifying information of any kind on the demographic sheet or questionnaires that would allow the researcher, or anyone else, to determine which person completed the materials.

Withdrawal Without Prejudice
Your participation is voluntary. You may choose not to take part at all. If you decide to participate, you are free to withdraw at any time. Leaving the study will not result in any penalty, and you will still receive the one (1) or two (2) research credits.

Cost of Participation
There will be no cost to you for participating in the current research study.

Alternative Procedures
Please see your class professor for any alternative procedures or assignments you can complete if you choose not to participate in this study.

Questions
If you have any questions regarding the research study or any possible research related injuries, please contact Robert J. Cramer at Crame001@bama.ua.edu or 348-5083, or Dr. Stanley Brodsky (faculty advisor for the current study) at sbrodsky@bama.ua.edu or 348-5083. If you have any questions about your rights as a research participant, you may contact Ms. Tanta Myles, The University of Alabama Research Compliance Officer, at 348-5152.

Debriefing Form

To the Participant:
If you have any questions or concerns following this session you may contact the primary investigator, Robert Cramer, at Crame001@bama.ua.edu. You may also contact the faculty supervisor, Stanley Brodsky, Ph.D., at SBrodsky@bama.ua.edu. If you have any questions about your rights as a participant, you may contact Ms. Tanta Myles, The University of Alabama Research Compliance Officer, at 348-5152.

The overall purpose of the current study is to assess the validity of the Witness Self-Efficacy Scale, a measure assessing one's ability to testify.

If you want to obtain the results of the study once data analysis is complete, you may email the primary investigator. He will keep your contact information on file and send the results once data has been analyzed.

Thank you for taking the time to participate in this study. Your cooperation is appreciated.

Sincerely,

Robert J. Cramer, M.A.
Department of Psychology
The University of Alabama
Gordon Palmer Hall 364C
Crame001@bama.ua.edu

APPENDIX D: WRITTEN ACCOUNT OF FALSE ACCUSATIONS

Instructions: Today you are serving as a criminal defendant wrongly accused of committing an act. As such, you will be required to testify about a time that you were falsely accused of something. Using the space below, please provide a written account about a time you were **falsely accused** of committing an action. Please note that it does not have to be a crime. Rather, it should be a less harmful action (e.g., lying, cheating on an exam, etc.). In order to comply with American Psychological Association Ethical Guidelines and University of Alabama Human Subjects Review Criteria, we ask that you avoid written accounts of any of the following: a) elder or child abuse or neglect, b) incidents containing excessive psychological or physical harm, or c) anything that could lead to a lawsuit or criminal charges. If you have questions about these, please ask the primary investigator, Robert Cramer.

APPENDIX E: CONTACT INFORMATION SHEET FOR MENTAL HEALTH RESOURCES

In order to comply with University of Alabama Human Subjects Review requirements, and ethical standards of the American Psychological Association, anyone scoring in the moderate to severe range on a measure of depression will receive this form. Below is contact information for university mental health resources. It is recommended that you seek assistance as soon as possible. If you are in need of immediate assistance, please inform the primary investigator, Robert Cramer, and he will aid you in seeking mental health assistance. Also, the research supervisor, Dr. Stanley L. Brodsky, is a licensed clinical psychologist and available for immediate consultation.

University of Alabama Psychological Clinic: (205) 348-5000
Gordon Palmer Hall

University of Alabama Counseling Center: (205) 348-3863
Russell Hall 225

University of Alabama Police Department: (205) 348-5454
800 Sorority Circle (New Hall)