

COMMUNICATING VIOLENCE RISK: JUDICIAL
DECISION MAKING IN INVOLUNTARY
CIVIL COMMITMENTS

by

STEPHANIE ANN EVANS

KAREN SALEKIN, COMMITTEE CHAIR

STANLEY BRODSKY
CARL CLEMENTS
LEE MALLORY
PATRICIA PARMELEE

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

2012

Copyright Stephanie Ann Evans 2012
ALL RIGHTS RESERVED

ABSTRACT

Beginning in the 1990s, researchers in the field of violence risk assessment recognized that even the most accurate and valid risk assessment could not assist fact-finders if violence risk was not communicated in a clear, precise, and complete manner (Monahan & Steadman, 1996; Schopp, 1996). Due to this growing attention to the importance of risk communication, four empirical studies have investigated how risk messages impact decisions and how decision-makers interpret risk messages (Kwartner, Lyons, & Boccaccini, 2006; Monahan et al., 2002; Slovic & Monahan, 1995; Slovic, Monahan, & MacGregor, 2000).

The main purpose of the current study was to investigate judges' opinions regarding the probative value of risk communication messages in civil commitment proceedings. There were five types of risk communication messages that were investigated in this study: (a) description model; (b) prediction model [categorical format]; (c) prediction model [probabilistic format]; (d) prediction model [frequency format]; and (e) management model. Secondary purposes of this study were to investigate whether these risk messages influence judicial decisions and whether the attributes of role orientation, legal philosophy, and Fear of False Negatives (FFN) impact the decision-making process.

A national sample of 403 judges completed the study, which produced an 18.28% response rate. Each participant was randomly selected to receive one of ten risk vignettes in which type of risk message and risk level were systematically varied. Participants answered three questions regarding the applicable risk vignette and completed demographic and judicial attribute questionnaires.

This study found that the risk models (i.e., description, prediction, and management messages) were viewed as equally probative. However, within the risk prediction model, categorical messages were rated as highest in probative value. Results indicated that risk prediction messages, in general, led to stricter rulings than did the other risk models, but no one risk prediction format (i.e., categorical, probabilistic, or frequency) resulted in higher restrictiveness in judicial ruling. Finally, this study found that legal philosophy and FFN, but not role orientation, impacted the restrictiveness of judicial rulings. Limitations of this study and recommendations for future research are discussed.

LIST OF ABBREVIATIONS AND SYMBOLS

FFN	Fear of False Negatives: A construct that refers to a judge's opinion regarding the consequences of an erroneous ruling
%	Percentage
r	Pearson product-moment correlation coefficient
=	Equal to
d	Cohen's
<	Less than
NCSC	National Center for State Courts
ANOVA	Analysis of Variance
USPS	United States Postal Service
p	Probability associated with the occurrence under the null hypothesis of a value as extreme as or more extreme than the observed value
Mdn	Median
SD	Standard deviation
t	Computed value of t test
α	Cronbach's index of internal consistency
χ^2	Computed value of a chi-square test
N	Total number in a sample
M	Mean: arithmetic average

F	Fisher's F ratio: A ratio of two variances
SE	Standard error (of measurement)
PPP	Positive predictive power
NPP	Negative predictive power
AUC	Area under the curve
η^2	Partial eta squared; measure of strength of relationship in an ANOVA
r^2	Coefficient of determination; refers to the adjusted squared correlation of an ANOVA
Z	A standard score
CI	Confidence interval
LL	Lower level
UL	Upper limit
β	Parameter estimate; log-odds (logit) regression coefficient
df	Degrees of freedom: number of values free to vary after certain restrictions have been placed on the data
-2LL	Minus 2 log-likelihood; used to assess the fit of a model by determining the probability the category of the dependent variable can be predicted from the values of the independent variables
R^2	Multiple correlation squared; refers to Nagelkerke's pseudo correlation squared estimate presented in a logistic regression model.

ACKNOWLEDGMENTS

Words cannot convey the happiness and excitement I feel in being so close to finishing graduate school. Beyond a doubt, this doctoral program has been a life-transforming journey. When reflecting on this process, I can identify several people that were instrumental in helping me reach this point. There are three groups of people I would like to acknowledge: those who helped me with completing this dissertation, those who helped me “enjoy” the process, and those who helped me believe I could accomplish this journey in the first place.

With regard to my dissertation, I would like to thank all of my committee members for their invaluable input and inspiring questions. I am most indebted to Dr. Karen Salekin for giving me the opportunity to be in such a great graduate program, and for reviewing this document. I will always remember our lab meetings with fondness. I am grateful to Dr. Lee Mallory, whose initial discussions of risk assessment and risk communication inspired the core of this dissertation and allowed me to brainstorm my ideas. Thanks to Dr. Patricia Parmelee for helping me develop this project during her grant writing class. Finally, to Drs. Carl Clements and Stanley Brodsky, I am eternally grateful for the knowledge I obtained from your classes and the insights you provided during prior dissertation meetings. To Debra Chen, I appreciate all your help with reviewing this document and getting it on the table. You are definitely one of the nicest people I have ever met. And of course, like most things in life, this project would not have been possible without money; as such, I am extremely thankful for the funding provided by an American Psychology Law Society Grant, an American Academy of Forensic Psychology Grant, and University of Alabama Research and Travel awards.

The second group, of people I would like to acknowledge, comprises of some of the strongest women in the world. I could not have accomplished this, or anything, without my closest friends. To my oldest (but not eldest) friend, Tricia Putnam: Life has taken us down some very different paths, but we always end up taking our little detours together and I will always be grateful for these moments of freedom. Prost, to our adventures in Europe, Asia, and America. I look forward to our future journeys. I love you, even if you are a Pisces. To my newest friend, Abbie Perelman: Thank you for making me laugh the last couple years and keeping all my (many) secrets. Praise Canada for being the catalyst of our friendship. Considering I will probably be on the phone with you in the next couple hours, I am sure we will have many future adventures together. I love you, even if you are a Cancer. To my dearest friend, Chozie (Nichols) Thorp: You are the most nurturing and supportive person I have ever met. I always feel so safe and unconditionally loved by you. It was hard to leave you to go to graduate school, but I am so impressed with how easy it is to slip back into our relationship whenever I see or talk to you. I love you and your babies (also, thank you for not being a water sign).

Finally, I am so appreciative of my family who gave me the strength and confidence to pursue my dreams. In many ways, this dissertation should be dedicated to my grandmother. She is the one who first put me on the path of psychology. As a child, she would discuss psychology and other intellectual topics. So to my grandmother Eunice Evans: Thank you for always talking to me like I was an adult; thank you for having an amazing book collection and letting me “borrow” your books for decades; and thank you for teaching me to think critically. You are the smartest person I have ever known. I love you and grandpa, who is the sweetest person I have ever known. To my daddy: Thank you for giving me so many opportunities in life. You always put my needs at the forefront. Most importantly though, you taught me that hard-work, perseverance,

and determination are the requirements to achieve anything worthwhile. I am eternally grateful for this lesson as these traits have allowed me to compensate for all my weaknesses. To my little brother: You are the most creative person I know. Thank you for those moments of reprieve, where I can discuss the important things in life, like my zombie apocalypse plan.

Para mi madre: Gracias por toda la fuerza que usted me ha dado. Tu siempre, y sin duda, ha creído en mi capacidad de lograr lo que sea (menos cantar). No hay palabras para expresar lo que siento por ti; sólo puedo decir que te quiero con todo mi corazón. Para mi Tía Miriam: Gracias por siempre rezar por mí y por ser tan buena persona. Para mi Tía Mercedes y mi abuela: Gracias por ser parte de mi vida. Para mi Tío Carlos: Su muerte todavía me duele, pero sé que tu estas siempre a mi lado, mi guía espiritual, y mi protector. Y finalmente, este triunfo no sera posible sin la Virgencita del Carmen: Gracias por ser mi luz y mi fuerza. Siempre voy apreciar la belleza, que es mi vida.

CONTENTS

ABSTRACT.....	ii
LIST OF ABBREVIATIONS AND SYMBOLS	iv
ACKNOWLEDGMENTS	vi
LIST OF TABLES.....	x
LIST OF FIGURES	xi
1. INTRODUCTION	1
2. METHODOLOGY	25
3. RESULTS	41
4. DISCUSSION.....	66
REFERENCES	80
APPENDIX.....	89

LIST OF TABLES

1. Risk Communication Models	16
2. Illustration of 5 (Type of Risk Message) X 2 (Risk Level) Design.....	25
3. Demographic Variables	42
4. Frequency of Probative Rating Endorsement by Risk Model	44
5. Frequency of Probative Rating Endorsement by Risk Format	46
6. Mean Difference, Standard Error, and Significance of Bonferroni Comparisons	46
7. Kruskal Wallis Statistics for Different Conceptualizations of Risk Format and Ruling	47
8. Classification Accuracy of Judicial Decisions by Prediction Format.....	49
9. Analysis of Variance for Risk Level predicting Probative Value by Risk Message	52
10. Dangerousness as a Mediator between Risk Format and Ruling varied by Risk Level	55
11. Dangerousness as a Mediator between Risk Model and Ruling varied by Risk Level	56
12. Classification Accuracy of Judicial Decisions by FFN Levels.....	58
13. Parameter Estimate Statistics for Interactions between Risk Level, FFN, and Risk Format...59	

LIST OF FIGURES

1. Conceptual Model Underlying Hypotheses.....	28
2. Mean of Probative Value varied by Risk Model and Risk Level.....	45
3. Mean of Probative Value varied by Risk Format and Risk Level.....	46
4. Impact of Risk Level on Probative Value varied by Risk Message.....	51
5. Distribution of FFN Scores.....	57
6. Distribution of Legal Philosophy Scores.....	60
7. Legal Philosophy as a Moderator between Dangerous Determination and Ruling.....	62
8. Distribution of Role Orientation Scores.....	64

CHAPTER 1

INTRODUCTION

In the United States, the process of civil commitment allows for “the state-sanctioned involuntary hospitalization of individuals with mental disorders” (Melton, Petrila, Poythress, & Slobogin, 2007, p. 325). Because the U.S. Constitution protects the liberty of American citizens from arbitrary and purposeless government actions, all legal actions, including involuntary civil commitment, must serve a purpose (Winick, 2005). Involuntary civil commitments were originally based on a medical model that served the *parens patriae* philosophy of the state (Mead & Nabors, 2006; Winick, 2005). *Parens patriae* is the belief that the state is responsible for the care of individuals who cannot care for themselves. Under this philosophy, hospitalization was thought to be in the respondent’s best interests and necessary to benefit his or her health. This philosophy resulted in the state deferring to the expertise of clinicians and the state allowing civil commitments to be a medical process instead of a legal process (Winick, 2005).

The *parens patriae* philosophy was challenged in the 1970s by a series of court cases (e.g., *Dixon v. Attorney General of the Commonwealth of Pennsylvania*, 1971; *Lessard v. Schmidt*, 1972; *Wyatt v. Stickney*, 1972) that revealed the deplorable conditions of psychiatric facilities or suggested that commitment was antitherapeutic. The Supreme Court ruling in *O’Connor v. Donaldson* (1975) held that civil commitment statutes were too broad and there was “no constitutional basis for confining such persons involuntarily if they are dangerous to no one” (p. 575). As a result of these cases, respondents could no longer be civilly committed

simply because they needed treatment; instead, a police power philosophy for civil commitments developed (Conroy & Murrie, 2007). Under the police power philosophy, the purpose of civil commitments is to protect the community from mentally ill respondents who may be dangerous to others (Mrad & Nabors, 2006; Winick, 2005). Currently, civil commitments have been referred to as an “uneasy mixture of the *parens patriae* and police powers” (Melton et al., 2007, p. 325).

Civil Commitment Criteria

As the purpose behind civil commitment has evolved, civil commitment statutes have incorporated criteria that reflect both *parens patriae* and police power philosophies. Although civil commitment statutes differ by state, all require the following two criteria for commitment: (a) the respondent must have a mental illness and (b) the respondent must be a danger to self or others as a result of the mental illness (Brooks, 2007; Melton et al., 2007; Parry, 1994; Schopp, 1996; Slovic & Monahan, 1995; Winick, 2005). The legal term *mental illness* should not be confused with the clinical term *mental disorder* because in many situations, the diagnosis of a mental disorder is not sufficient to fulfill the requirement of mental illness (American Psychiatric Association, 2000). Instead, mental illness requires that there exists current symptoms that impede an individual’s ability to function on a day to day basis. In most legal statutes, mental illness is a “disorder of thought, mood, perception, orientation, or memory” that “impairs judgment, behavior, capacity to recognize reality, or ability to meet the ordinary demands of life” (Melton et al., 2007, p. 335). In order to describe the extent to which the mental illness impairs the respondent, statutes typically use terms such as “significant, severe, substantial, or gross” (Winick, 2005, p. 48) to describe the impairment that results from the mental illness (Melton et al., 2007).

The majority of statutes use the phrase “danger to self or others” as their second criterion. It is important to note that this criterion contains two separate concepts. The first concept, danger to self, refers to the possibility the respondent is in jeopardy of harming him- or herself by not meeting the basic needs of food, shelter, hygiene, and security or by engaging in intentional suicidal or self-injurious behavior (Conroy & Murrie, 2007; Melton et al., 2007; Winick, 2005). Danger to self reflects the *parens patriae* philosophy because it requires that the respondent’s decision-making skills, and awareness of the need for treatment, be so impaired that hospitalization is necessary to prevent life-threatening deterioration (Melton et al., 2007; Winick, 2005).

The second concept, danger to others (also referred to as dangerousness), reflects the police power goal of protecting the public from dangerous individuals (Winick, 2005). Although criminal courts exist to protect the public, the dangerous actions of a mentally ill respondent are conceptualized as directly resulting from his or her mental illness (Winick, 2005). Thus, criminal proceedings are considered inappropriate for mentally ill respondents, because their actions are not considered deserving of punishment. Although criminal courts and civil commitment proceedings can share the goal of incapacitation, civil commitment does not serve the objective of punishing the respondent (Melton et al., 2007; Winick, 2005).

In addition to mental illness and dangerousness, some civil commitment statutes cite supplementary criteria for commitment. For example, some statutes consider, as separate criteria, a respondent’s incapacity to make treatment decisions and need for treatment. In other statutes, these two criteria are encompassed in the mental illness criterion (Melton et al., 2007). The criterion inability to care for self or grave disability is usually subsumed into the criteria of mental illness and/or danger to self. Finally, many statutes mention the requirement that

respondents receive the least restriction of their civil liberties necessary to accomplish the interests of the state (Melton et al., 2007; Winick, 2005).

Due to the requirement for the least restriction necessary, many civil commitment statutes allow judges to recommend that a respondent receive outpatient psychiatric treatment, as opposed to inpatient psychiatric treatment (Melton et al., 2007). The criteria for outpatient commitment tend to be similar to, or identical to, the criteria for inpatient commitment (Melton et al., 2007; Swartz, Swanson, Kim, & Pertila, 2006). Even in states without an explicit outpatient commitment option, outpatient treatment can still be considered during civil commitment proceedings (Swartz et al., 2006). Respondents court-ordered to outpatient treatment must adhere to a community treatment plan that may include medication, individual and/or group therapy, and supervision of living arrangements (Allbright, Levy, & Wagle, 2002). If the respondent fails to comply with outpatient treatment, the need for inpatient commitment will be reassessed by the judge (Swartz et al., 2006).

Judicial Decision Making

In American civil and criminal cases, the trier-of-fact has the obligation to determine whether the criteria applicable to the legal proceedings have been sufficiently met. According to Schopp (1996, 2001), judges¹ make two types of civil commitment decisions. First, the judge must decide if the criteria for commitment are present. This decision is referred to as the descriptive component. The second decision, referred to as the normative component, involves the determination of whether the strength of the criteria is sufficient to warrant involuntary civil commitment.

¹ It is acknowledged that in some jurisdictions, the judicial officer who makes the commitment decision may be called a magistrate, commissioner, or probate judge. For the sake of clarity, the term “judge” is used throughout this paper to refer to any judicial official who oversees civil commitment proceedings.

With regard to the dangerousness criterion, the descriptive component is established if the respondent presents a “pattern of circumstances and conduct” indicative of a “risk to harm” (Schopp, 2001, p. 219). The normative component addresses whether this risk is sufficient to necessitate involuntary civil commitment (Schopp, 2001). When judges make normative risk decisions, they are supposed to take into account five factors: (1) the type of violent behavior, (2) the extent of harm due to the violent behavior, (3) the likelihood of the violent behavior, (3) the imminence of the violent behavior, and (5) the frequency of the violent behavior (Heilbrun, O’Neill, Strohmman, Bowman, & Philipson, 2000; Melton et al., 2007). Although judges are supposed to consider all the aforementioned factors, researchers and scholars have suggested that judges tend to focus primarily on the likelihood of the violent behavior (Heilbrun et al., 2000; Litwark, 2002; Schopp, 1996). Because the U.S. Supreme Court has not set a threshold for likelihood of dangerousness that justifies commitment, judges must make normative decisions by identifying their own personal decision threshold (i.e., the specific subjective point in which the probability is large enough to be considered sufficient; Winick, 2005).

To date, only one study has examined the threshold at which dangerousness is considered sufficient to warrant commitment. Monahan and Silver (2003) asked 26 judges what probability of future violent behavior (i.e., 1%, 8%, 26%, 56%, and 76%) is sufficient to justify civil commitment. In this study, judges endorsed a conservative threshold with 61.6% of them deeming less than a 26% likelihood of violent behavior to be sufficient to necessitate commitment.

There appear to be only two other studies that have investigated decision making in civil commitment proceedings. Bursztajn, Gutheil, Mills, Hamm, and Brodsky (1986) asked five Massachusetts district court judges to rate the impact that 26 factors had on 41 of their civil

commitment decisions. In this study, judges reported that the following three factors had the strongest impact on their decisions: (1) testimony of the clinician, (2) whether the respondent would be a reliable outpatient, and (3) whether the respondent was able to take care of him- or herself. The authors pointed out these factors are notions emphasized by the *parens patriae* philosophy.

Bursztajn, Hamm, and Gutheil (1997) asked one Colorado probate court judge to rate the impact 26 factors had on 27 of his civil commitment decisions. In addition, the authors analyzed the extent to which each factor predicted commitment decisions. The authors concluded the results of this study were similar to that of Bursztajn and colleagues (1986), in that *parens patriae* considerations had a greater impact on the judge's decisions than did police power considerations. The authors based their conclusion on the finding that the three factors that had the strongest, statistical, impact on the judge's decision were *parens patriae* notions: (1) whether the family or friends of the respondent favored commitment, (2) whether the respondent would be a reliable outpatient, and (3) whether the respondent seemed unable to care for him- or herself.

Predictors of Judicial Decisions

There are two major schools of thought regarding which factors a judge considers when making a legal decision (Brisbin, 1996; Gillman, 2001; Wrightsman, 1999). During the early 20th century, the predominant view of legal scholars (e.g., Corwin, 1924; Levi, 1948) was that judges make their decisions by considering only legal factors. According to Wrightsman (1999), the fundamental stance of this position was that judges only considered “the facts and issues of the current case and relate[d] them to previous decisions, to applicable laws, and to the Constitution” (p. 19). Although some subsequent legal scholars have also maintained this stance (e.g.,

Mendelson, 1963; Schauer, 1988), the predominance of this view decreased due to research that suggested that decision making is a complex process (Gillman, 2001; Wrightsman, 1999).

Instead of the view that judges only consider legal factors, legal scholars (e.g., Frank, 1949; Llewellyn, 1930; Rowland & Carp, 1996; Segal & Spaeth, 1996) have postulated that extralegal factors influence judicial decisions. For example, research has indicated the following extralegal factors may influence judicial decisions: (1) a judge's opinion about his or her role in the judiciary, (2) a judge's opinion about the purpose of his or her ruling, and (3) a judge's opinion about the consequences of his or her errors (Gibson, 1977, 1978; Guthrie, Rachlinki, & Wistrich, 2001; Homel & Lawrence, 1992).

In order to better understand the judicial decision-making process, researchers have recommended that the mediating or moderating role of extralegal factors be investigated (Gibson, 1978, 1983; Homel & Lawrence, 1992). Instead of examining the direct impact these factors have on judicial outcomes, research has focused on whether these factors can explain how/why (i.e., mediating role) or when (i.e., moderating role) a specific judicial ruling will occur. For example, a defendant's postconviction statement may impact the judge's perception of the defendant's remorse, which in turn may lead to a more lenient sentence (Slovenko, 2006; Ward, 2006). This is an example of a mediating relation because it explains how/why a decision is made; in this case, the lenient sentence occurred because the statement of remorse decreased the punitive focus of the judge. In contrast to a mediating relation, a moderating relation occurs when a moderator affects the extent to which one variable (i.e., the independent variable or predictor) affects another variable (i.e., the dependent variable). For example, if a judge believes most defendants are dishonest, then a postconviction statement may have no impact on the sentence. In this case, the judge's preexisting belief about the veracity of postconviction

statements (i.e., moderator) affects the strength between the statement of remorse (i.e., predictor) and sentence recommendation (i.e., dependent variable).

The aforementioned recommendation to examine the mediating and moderating role of extralegal factors is in line with a cognitive approach to understanding decision making. The cognitive approach suggests that people process details about a decision through preexisting schemas (Kahneman & Tversky, 1973). Schemas are “cognitive structure[s] that represent knowledge about a concept or type of stimulus, including its attributes and the relations among the attributes” (Fiske & Taylor, 1991, p. 98). Schemas allow the mind to process information quickly because knowledge is stored at inclusive levels (i.e., the overall pattern or category) rather than every original experience being stored in its raw form (Fiske & Taylor, 1991). In order to process information at inclusive levels, individuals utilize shortcuts (also called heuristics) that allow them to classify new information into its appropriate category (i.e., schema; Fiske & Taylor, 1991). Although schemas help information be processed in a quick and efficient manner, heuristics may cause an individual to ignore or minimize relevant evidence and focus on irrelevant factors (Anderson, Pichert, & Shirey, 1983; Kahneman & Tversky, 1973).

With regard to judicial decision making, some researchers have found that judges rely on heuristics (i.e., a strategy or principle that allows one to make rapid estimations or predictions) in order to create schemas, which result in the acceleration of the decision-making process (Fariña, Arce, & Novo, 2003; Guthrie et al., 2001; Kulik, Perry, & Pepper, 2003; Viscusi, 1999; Von Helversen & Rieskamp, 2009; Wrightsman, 1999). For example, some judges engage in the representative heuristic, in which they ignore important background statistical information and favor vivid, individuating evidence (Guthrie et al., 2001). A hypothetical example of the use of this heuristic would be a judge who ignores research that indicates sexual offenders have a

recidivism rate of 15%, and instead creates a schema that all sexual offenders will recidivate based on memories of vivid news programs that featured sexual offenders that recidivated. The possibility that judges utilize schemas is further supported by research that has investigated the constructs of role orientation and legal philosophy (e.g., Gibson, 1978; Homel & Lawrence, 1992), and the scholarly conjectures regarding the construct of fear of false negatives (FFN; e.g., Wexler, 1983; Winick, 2005). These three constructs will be discussed in the sections that follow.

Role orientation.

Role orientation is “a belief about proper behavior within an institutional position” (Gibson, 1983, p. 17). In particular, this construct usually refers to the extent to which judicial decisions adhere to statutes and precedents (Baum, 1997; Gibson, 1978; 1983; Van Koppen & Kate, 1984). Most of the research on this construct has been devoted to the construction of typologies or systematic classifications (e.g., Blanck, Rosenthal, Hart, & Bernieri, 1989; Gibson, 1983; Howard, 1977; Scheb & Unga, 1986; Unga & Baas, 1972; Wold, 1974). Instead of conceptualizing role orientation as being on a spectrum, *systematic classification* research results in the creation of distinct categories of role orientation (e.g., law interpreters; activists) with accompanying characteristics that make up this category or type of judge. For example, judges who strictly adhere to statutes and precedents were deemed to be “law interpreters,” while judges who subordinate statutes and precedents were referred to as “lawmakers” or “activists” (Gibson, 1983; Unga & Baas, 1972).

These categorical conceptualizations of role orientation have been used in research that has found that this construct has a moderating impact on judicial outcomes (Gibson, 1978, 1983; Scheb, Bowen, & Anderson, 1991; Scheb, Unga, & Hayes, 1989; Simons, 1998; Smith, 1993).

For example, in one study, judges' attitudes toward the criminal justice system and political issues explained 14% of the variance in judicial ruling, while the interaction between role orientation and these attitudes explained 64% of the variance (Gibson, 1978). Similarly in another study, the correlation between political ideology and judicial ruling was .63 for judges considered to be activists and .2 for judges who strictly follow legal rules and precedents (Scheb et al., 1991); $r = .63$ is considered a strong correlation, whereas $r = .2$ is considered a weak correlation (DeCoster, 2007). In conclusion, role orientation appears to be related to the extent to which judges rely on extralegal factors when they make judicial decisions (Gibson, 1977, 1978; Scheb et al., 1989).

Legal philosophy.

Legal philosophy refers to a judge's perceived objective for legal proceedings. Within the context of criminal cases, five sentencing goals or penal philosophies have been identified: (1) punishment, (2) rehabilitation, (3) incapacitation, (4) general deterrence, and (5) specific deterrence (Carroll, Perkowitz, Lurigio, & Weaver, 1987). To date, research has demonstrated a judge's goal preference has a significant impact on his or her ruling (Clancy, Bartolomeo, Richardson, & Welford, 1981; Davis, Severly, Kraus, & Whitaker, 1993; Homel & Lawrence, 1992). For example, one study found that federal district judges' perceptions of sentencing goals accounted for 23% ($r = .48$) of the variance in their decisions regarding the length of prison sentence (Clancy et al., 1981). Homel and Lawrence (1992) found that the goal of sentencing affected which case details the judge relied on when they made judicial decisions. Specifically, judges with the goal of "protecting the public" (i.e., incapacitation) were more influenced by the number of prior offenses and type of legal representation than were judges that aligned with the goal of deterrence.

Although this construct has not been examined with regard to civil commitments, Hiday (1983) opined that a judge's opinion regarding the purpose of commitment could predict judicial outcome. For example, Hiday (1983) postulated that judges who view commitment as punitive and endorse a police power philosophy may be less inclined to commit a respondent. Judges who believe that commitment is beneficial and endorse a *parens patriae* stance may be more inclined to commit a respondent (Hiday, 1983).

Fear of false negatives (FFN).

FFN refers to a judge's perception about the consequences of an erroneous ruling. Scholars (e.g., Durham, 1996; Wexler, 1983; Winick, 2005) have suggested that some judges err on the side of caution when they make commitment decisions. One reason why judges may over predict the need for commitment is that a judge may receive negative feedback if he or she commits a false negative error (Wexler, 1983). A false negative occurs when a respondent who is deemed not dangerous enough to justify commitment commits a violent act shortly after the decision was made. The feedback to the judge may depend on the seriousness of the violent behavior because some behaviors (e.g., destruction of property; minor slapping; pushing of others) are less likely to be reported to the court (Durham, 1996). In contrast, serious violent behaviors (e.g., murder; assault resulting in injury) are very likely to come to the attention of the court and in some cases attract the attention of the media and community in general (Durham, 1996; Wexler, 1983). In these scenarios, the credibility of the decision maker may be questioned and he or she may be judged harshly by community members (Mulvey & Lidz, 1985; Wexler, 1983, 1992; Winick, 2005). A judge is unlikely to incur negative feedback from a false positive error (i.e., committing a respondent who would not be dangerous), because most of the time, false positives will not come to the attention of the decision maker or community members.

Instead, the respondent's lack of violent behavior will likely be attributed to the restrictive course of hospitalization and/or the benefits of the treatment (Wexler, 1983).

Even though judges are unlikely to incur negative feedback from a false positive error, some judges may believe false positives are more problematic than false negatives. As noted in *Addington v. Texas* (1979) and *Goetz v. Crosson* (1992), some judges believe that the harm of erroneously committing an individual is equal to, if not greater, than the harm that results from not committing someone who needs it. In *Addington v. Texas* (1979, p. 427), the Supreme Court ruled that “the individual should not be asked to share equally with society the risk of error when the possible injury to the individual is significantly greater than any possible harm to the state.” That is, the justices in *Addington v. Texas* (1979) opined that the unconstitutional restriction of a respondent's civil liberties is significantly worse than any possible harm that may occur to the state. In addition, Appelbaum (1992) pointed out that false positives are costly because of the financial burden they place on psychiatric facilities.

Clinicians' Role in Civil Commitments

Commitment decisions are ultimately at the sole discretion of the judge. However, clinicians' testimony has been found to have an impact on judicial outcome (Winick, 2005). A review of the literature indicated that, to date, only two studies (Bursztajn et al., 1986, 1997) have investigated judges' opinions regarding the importance of expert testimony in civil commitment proceedings; both studies found that judges rate expert testimony as the most important factor in their commitment decisions. These findings are supported by the Supreme Court decisions in *O'Connor v. Donaldson* (1975) and *Addington v. Texas* (1979), in which the justices opined that clinicians contribute important information to civil commitment proceedings.

When clinicians participate in commitment proceedings, they usually gather information about “a person’s mental illness, treatment needs, dangerousness, and ability to survive outside of the hospital or in some less restrictive alternative” (Melton et al., 2007, p. 354). The mental illness, treatment needs, and least restrictive alternative aspects of the evaluation require the clinician to examine clinical diagnoses and to determine how treatment could lessen the impairment caused by the mental illness. The dangerousness aspect is addressed by a clinical assessment of suicide risk and/or violence risk. A suicide risk assessment requires the identification of factors associated with suicide or self-injurious behavior (e.g., suicidal plans; access to weapons; history of prior suicide attempts). Similarly, violence risk assessments require the identification of risk factors associated with violent behavior toward others (e.g., past violence; substance abuse; young age). According to Winick (2005, p. 63), “Civil commitment courts typically rely upon the testimony of clinical expert witnesses who have evaluated the individual and who present their clinical conclusions concerning the degree of risk he or she is thought to present.”

Violence Risk Assessments in Civil Commitments

Risk assessments for civil commitment proceedings differ from those in other forensic contexts (Melton et al., 2007). In civil commitment proceedings, risk assessments must be completed within a short time frame, usually within a period of 72 hours (Melton et al., 2007). This hastened pace, generally, does not allow clinicians to obtain the records and collateral sources typically utilized in criminal risk assessments. As such, time limitations may result in less thorough assessments (Melton et al., 2007).

Despite these limitations, Melton and colleagues (2007) recommend clinicians conduct risk assessments that identify empirically-validated risk factors known to increase a respondent’s

likelihood of committing a violent act. Research has indicated that the following two risk factors are most predictive of violence risk in psychiatric populations: (1) recent substance use (Douglas & Skeem, 2005; Monahan et al., 2001; Steadman et al., 2000; Volavka & Tardiff, 1999); and (2) history of past violence (Conroy & Murrie, 2007; Monahan, 1981; Monahan et al., 2001; Scott & Resnick, 2006; Steadman et al., 2000). Other risk factors that have obtained empirical support include medication noncompliance (Bartels, Drake, Wallach, & Freeman, 1991; Monahan et al., 2001; Swanson et al., 2000), young age (Scott & Resnick, 2006; Swanson, Holzer, Ganju, & Jono, 1990), and recent unemployment (Steadman et al., 2000). Risk factors can be either static (i.e., unable to be changed) or dynamic (i.e., modifiable).

Although consideration of risk factors is crucial for a valid and reliable risk assessment, clinicians are also advised to consider protective factors in order to communicate more fair and balanced risk findings (Rogers, 2000). Protective factors are processes that mitigate the expression of maladaptive behavior (Lodewijks, Ruiters, & Doreleijer, 2010; Werner, 1995). Research with juvenile samples has found that prosocial involvement, positive role models, and familial support are just a few examples of protective factors (e.g., Herrenkohl et al., 2003; Lodewijks et al., 2010; Morrison, Robertson, Laurie, & Kelly, 2002; Vance, Bowen, Fernandez, & Thompson, 2002; Werner, 1995).

Risk Communication

Beginning in the 1990s, researchers in the field of violence risk assessment recognized that even the most accurate and valid risk assessment could not assist fact-finders if violence risk was not communicated in a clear, precise, and complete manner (Monahan et al., 2002; Schopp, 1996). For example, in 1996 a series of theoretical articles published in the *American Psychologist* emphasized the need to improve risk communication (e.g., Grisso & Tomkins,

1996; Monahan & Steadman, 1996; Schopp, 1996). Since that time, scholars (e.g., Monahan & Steadman, 1996; Schopp, 1996) have debated the probative value and ethical repercussions of risk communication messages. Research into clinicians' use of and perceived value of risk messages has also been conducted (i.e., Heilbrun et al., 2000, 2004; Heilbrun, Philipson, Berman & Warren, 1999).

Risk models.

Three models of risk assessment communication have been identified: (1) risk description model, (2) risk prediction model, and (3) risk management model (Borum & Verhaagen, 2006; Heilbrun, 1997). The risk description model involves the identification and description of relevant nomothetic (i.e., general or universal findings) or idiographic (i.e., individualized, case specific findings) risk and protective factors (Conroy & Murrie, 2007). Individuals who use the description model do not offer a prediction of future violence, nor do they offer recommendations for intervention strategies (e.g., drug counseling to ameliorate substance abuse; long-acting injectable medication to ameliorate medication noncompliance). Conclusive statements regarding predictions or recommendations are excluded because the purpose of this model is to simply inform decision makers of the risk and protective factors relevant to a particular case.

In contrast to the risk description model, the risk prediction model includes explicit conclusions regarding an individual's likelihood to commit a violent act in the near future. This conclusion is provided after, and based on, a description of the relevant risk and protective factors. The primary goal of the risk prediction model is "to attempt to determine, as accurately as possible, the probability of a specified event's occurrence" (Heilbrun, 1997, p. 352). This

model does not include recommendations regarding strategies or interventions that may reduce the likelihood of the predicted event.

The risk management model moves beyond the description and prediction models by offering specific strategies or interventions to reduce the impact of identified dynamic risk factors (Conroy & Murrie, 2007; Heilbrun, 1997). The risk management model is considered the most comprehensive model because it encompasses elements from the description model (i.e., identification of risk/protective factors) and, occasionally, the prediction model (i.e., prediction statement; Conroy & Murrie, 2007). The main emphasis of the risk management model is on risk reduction and thus, recommendations for risk-reducing interventions or strategies are considered the most important element of this model (see Table 1 for summary of the three risk models).

Table 1

Risk Communication Models

Risk Model	Identifies Risk/Protective Factors	Conclusion Statements	Goal or Purpose
Risk Description	Yes	None	Inform
Risk Prediction	Yes	Prediction of future risk	Predict behavior
Risk Management	Yes	Intervention strategies (may also include prediction of risk)	Reduce risk

In addition to differences regarding conclusive statements and goals, these three risk models differ with regard to the extent that they are valued by clinicians. Two studies have investigated the perceived probative value of these models (Heilbrun et al., 2000, 2004). Heilbrun and colleagues (2000) asked 71 experts in the field of risk assessment (i.e., 41 psychologists, 28 psychiatrists and 2 sociologists) to read eight risk vignettes and for each vignette to rate the probative value of risk prediction, risk management, and risk description messages. To clarify, all eight risk vignettes were presented to each participant and the vignettes

differed with regard to risk level (high versus low), risk factors (static versus dynamic), and legal context (commitment versus parole decision); all the risk messages were presented with each vignette. The results indicated that the risk management model was the most highly valued form of risk communication, especially in high risk scenarios. However, risk experts who primarily conduct research rated the risk management messages as less probative than did risk experts who are involved in clinical practice. In a subsequent study that included a sample of 256 members of American Psychological Association (APA), who reported interest in either clinical or forensic psychology, Heilbrun and colleagues (2004) also found that the risk management model was rated as most probative in value.

Although the aforementioned research has provided tentative support that clinicians find risk management messages to be probative when they make commitment decisions, there is a lack of consensus regarding which model is ethically appropriate to use in civil commitment proceedings. For example, Schopp (2001) stated the risk description model is the only appropriate model for clinicians to use. He pointed out some civil commitment statutes (e.g., Nebraska; Wisconsin) indicate that the temporal focus of the dangerousness criterion is the present and not the future. According to Schopp (2001) the referral question does not require a prediction and clinicians who use risk prediction messages “distort the meaning of ‘dangerousness’ in the statute” (p. 221). In contrast to Schopp’s (2001) stance, Heilbrun (1997) stated that the initial stage of involuntary civil commitments require the use of the risk prediction model to communicate findings. The temporal focus of dangerousness is interpreted as the imminent future and a prediction is considered informative for judges forming a decision about whether the risk is sufficient to warrant commitment. With regard to the risk management model, Schopp (2001) indicated the use of this model was inappropriate in civil commitment

proceedings because it results in the clinician having to go beyond the referral question. To date, no studies or theoretical articles have examined judges' opinions regarding the appropriateness of risk models in civil commitment proceedings.

Risk prediction formats.

As previously mentioned, risk prediction messages include explicit predictions of the individual's likelihood to commit a violent act in the near future. This prediction can be communicated through three different formats. Specifically, clinicians use either categorical or numerical (i.e., probabilistic or frequency) formats to communicate their results. Categorical risk estimates usually follow a simple ordinal scale of low, moderate, or high risk (Conroy & Murrie, 2007; Heilbrun et al., 1999). If a clinician chooses a numerical format, he or she may decide to communicate their risk message through probabilistic statements (e.g., 76% chance the respondent will commit a violent act) or through frequency statements (e.g., 76 out of 100 people with similar features as the respondent will commit a violent act).

Similar to the debate about the appropriateness of the risk models, there is a lack of consensus regarding which risk prediction format to use in civil commitment proceedings. Originally, numerical formats, in particular probabilities, were considered the ideal way to communicate risk predictions (e.g., Monahan & Wexler, 1978; Steadman, 1987; Steadman et al., 1994). The rationale was that as the use of actuarial measures increased, risk estimates would become more valid and more likely to be expressed in probabilistic terms (Monahan & Wexler, 1978; Steadman et al., 1994).

Although Monahan and Steadman were initially in favor of only using probabilistic estimates (e.g., Monahan & Wexler, 1978; Steadman, 1987; Steadman et al., 1994), in their 1996 article they changed their stance. Drawing from research that examined how to communicate the

risk of meteorological events, Monahan and Steadman (1996) postulated that decision makers may be unable to accurately process risk probabilistic estimates. The authors cited the findings of Baker (1995), which suggested that individuals interpret probabilistic estimates of rare events in comparative or ordinal ways. To clarify, Baker (1995) found that people are more likely to evacuate if the probability of a hurricane hitting their area was noticeably higher than the probability for a neighboring area, regardless of the actual estimates used. The author concluded probabilistic estimates of violence may be vulnerable to similar misinterpretation because violent acts are considered rare events. Due to the potential for misinterpretation, Monahan and Steadman (1996) recommended that categorical messages be used in combination with probability or frequency messages (e.g., “this patient is considered a low risk, which means that this patient is believed to have a 29% likelihood of being violent”).

Heilbrun and colleagues' (1999) study provided support for Monahan and Steadman's (1996) belief that probabilistic estimates may be misinterpreted. Specifically, Heilbrun and colleagues (1999) found that 25% of clinicians opined that probability messages are easily misinterpreted. In fact, only 2% of clinicians reported that they use numerical formats when they communicate risk. The majority of clinicians (49%) stated they avoided probabilistic estimates because probabilities give the impression that clinicians are more confident in their predictions than is warranted. Hilton and colleagues (2005) have since postulated that as actuarial measures improve, clinicians will become more comfortable in their use of numerical estimates. Hence, these authors suggested that the findings of Heilbrun and colleagues (1999) may not be valid in the future.

Perception of Risk Prediction Messages

In addition to the scholarly speculation (i.e., Heilbrun et al., 1999; Hilton et al., 2005; Monahan & Steadman, 1996) regarding how decision makers interpret risk prediction messages, researchers have investigated how risk messages impact decisions (i.e., Kwartner, Lyons, & Boccaccini, 2006; Monahan et al., 2002; Slovic & Monahan, 1995; Slovic, Monahan, & MacGregor, 2000). Specifically, three empirical studies have examined the impact of risk prediction messages (i.e., categorical, probability, and frequency) on clinicians' decisions about dangerousness or discharge recommendations (Monahan et al., 2002; Slovic & Monahan, 1995; Slovic et al., 2000); one empirical study has investigated which risk prediction format judges find most useful in civil commitment proceedings (Kwartner et al., 2006). Studies have not yet been conducted on the impact that risk description or risk management messages have on clinicians' or judges' decisions, nor has any study investigated judges' perceived value of description or management messages.

The first study to investigate the impact of risk prediction messages was conducted by Slovic and Monahan in 1995. The authors asked 137 forensic clinicians (i.e., 93 psychologists, 18 psychiatrists, and 15 social workers) to rate the probability that eight different, hypothetical respondents would commit a violent behavior during the three years following the examination. In addition, participants were asked to make dichotomous determinations of whether each respondent was "dangerous." Results indicated that assigned probability was strongly dependent on the form of response scale the participants were given. Specifically, participants assigned lower risk probabilities to the respondent when given a response scale that included "smaller probabilities" (i.e., less than .001% to greater than 40%), than when given a response scale that included "larger probabilities" (i.e., 0% to 100%). However, the response scale did not affect

determinations of dangerousness. The discrepant results, that response scales affected probabilistic estimates but not dichotomous decisions, led the authors to conclude that clinicians use probabilities in a comparative or ordinal way without maintaining a consistent view of the actual values.

Slovic and colleagues (2000) continued to investigate the impact of risk prediction messages with a sample of 470 forensic psychologists and 409 forensic psychiatrists, who were asked to rate the likelihood that six different respondents would commit a violent act; the scenarios for this study were based on real discharge summaries. Participants were asked to provide predictions based on either frequency or probabilistic response scales and were asked to determine whether each respondent presented a low, medium, or high risk of harming someone after discharge. Similar to the authors' prior results, a clinician's prediction was dependent on the nature of the response scale. Specifically, participants assigned lower risk estimates when given small frequency response scales than when given larger frequency response scales. In addition, clinicians rated a respondent as posing a higher risk for future violence when given the information in a frequency statement, as compared to a probabilistic statement, even when the two formats were describing identical levels of risk (i.e., 20% versus 20 out of 100). The authors suggested that these findings may be due to clinicians' ability to visualize frequencies, or "imag[ine] the numerator" (p. 285), more easily than to visualize probabilities. To clarify, individuals have a tendency to focus on the first number (i.e., numerator) in a frequency statement and disregard the next number (i.e., denominator); this does not occur with probabilities because there is only one number (Yamagshi, 1997).

Monahan and colleagues (2002) examined the visualization hypothesis by asking 324 forensic clinicians and 466 nonforensic clinicians to determine whether they would discharge a

hypothetical respondent “from a hospital” (p. 122). The risk vignettes were systematically varied by message format (i.e., probabilistic or frequency) and outcome description (i.e., vivid or pallid description). Forensic clinicians were significantly more likely to decline release when given the frequency message, as opposed to the probability message. This result was not found for participants who worked in nonforensic settings. Similarly, forensic clinicians were more likely to decline release when given a vivid outcome as opposed to the pallid outcome. The vivid description used in this study was: “Recently, another patient who was discharged from the same facility killed a stranger in the community by smashing her skull with a baseball bat, resulting in her instant death” (p. 122). In contrast, the pallid description was “Recently, a stranger in the community sustained fatal injuries from another patient who was discharged from the same facility” (p. 112). Again, this result was not found for nonforensic clinicians. The authors concluded clinicians who work in forensic settings may be sensitized to visualizing frequencies because these clinicians may have more “personal experiences with making false negative predictions and the disastrous effects those predictions can have for the victims and for one’s own career” (p.126). Furthermore, the authors postulated that frequency messages may be “fear-inducing risk communication formats” (p. 126).

Although the aforementioned research is beneficial for the understanding of risk communication, research into judges’ opinions regarding risk prediction messages is more informative because they are the “consumers of risk communication” (Conroy & Murrie 2007, p. 112). A study by Kwartner and colleagues (2006) has been cited as the only investigation on this topic (Conroy & Murrie 2007). Kwartner and colleagues (2006) asked 116 judges to rate the probative value of probability (e.g., “76%”), frequency (e.g., “76 out of every 100”) and categorical (e.g., “high risk”) risk prediction messages. Results indicated that judges placed

significantly higher probative value on categorical messages than probability ($d = .22, p < .05$) or frequency messages ($d = .29, p < .05$). Also, results indicated that judges find any risk prediction message to be more probative in high risk scenarios than in low risk scenarios. Although Kwartner and colleagues (2006) provided valuable initial information on judges' perceptions, the authors acknowledged that the generalizability of their findings was "limited by the low response rate [12.5%] and the fact that judges were sampled from three southern states" (p. 193).

Purpose

The main purpose of the current study was to investigate judges' opinions regarding the probative value of different types of risk messages. This information was considered valuable because there is a lack of consensus and lack of research regarding which risk messages are appropriate for use in civil commitment proceedings. By expanding on the findings of Kwartner and colleagues (2006), the current study provided information regarding which type of risk message judges considered most probative in commitment proceedings. In addition, the current study investigated the extent to which risk messages influenced judicial decisions or the decision-making process. Clinicians may use the results of this study to guide their decisions regarding which risk messages to use when they communicate their results.

The current study examined the topic of risk communication because researchers (Heilbrun, Dvoskin, Hart & McNiel, 1999; Heilbrun et al., 2000, 2004; Kwartner et al., 2006; Monahan et al., 2002) have identified various reasons that the investigation of risk communication is of theoretical and practical importance. For example, risk communication is considered important because it is a concept designed to protect individuals from harm. Without risk communication, there is no link between risk assessments and decisions regarding risk.

Better-informed risk decisions can best be achieved when violence risk is communicated in a manner in which it can be clearly understood by the decision maker.

Although the primary focus of the current study was on risk communication, the current study also sought to contribute to the judicial decision-making literature. Specifically, one of the goals of this study was to obtain a better understanding of how judicial attributes (i.e., role orientation, legal philosophy, and FFN) impact the decision-making process. As Viscusi (1999) stated, in order to help control “risk sensibly,” researchers “need to understand whether decisions are flawed and, if so, in what way” (p. 27). As previously described, flawed decision making in civil commitment proceedings has negative consequences for the respondent (e.g., not receiving the appropriate level of care) and/or the community (e.g., being a victim of the respondent’s violent act; financial burden stemming from unnecessary hospitalization). Through an investigation into the impact of judicial attributes, the results of the current study provide empirical information regarding the extent to which extralegal factors impact civil commitment decisions.

CHAPTER 2
METHODOLOGY

Design

The current study examined the perceived value of different forms of risk messages through a national survey of judges who have jurisdiction over involuntary civil commitments. Participants were randomly selected to receive one of ten risk vignettes in which two factors were systematically varied in a 5 x 2 between-groups design (see Table 2). The dependent variables were probative value and judicial ruling. Dangerousness determination was a mediator variable between type of risk message and judicial ruling.

Table 2

Illustration of 5 (Type of Risk Message) X 2 (Risk Level) Design

Type of Risk Message	Risk level	
	(1) High risk	(2) Low risk
(1) Description	Condition One	Condition Six
(2) Prediction using categorical terms	Condition Two	Condition Seven
(3) Prediction using frequency terms	Condition Three	Condition Eight
(4) Prediction using probabilistic terms	Condition Four	Condition Nine
(5) Management	Condition Five	Condition Ten

In addition to the examination of risk communication, this study investigated the moderating role of the following constructs: FFN, role orientation, and legal philosophy. In particular, this study examined whether role orientation or legal philosophy moderated the relation between dangerousness determination and judicial ruling. This study also investigated

whether FFN had a direct impact on judicial ruling and whether FFN moderated the relation between risk messages and judicial ruling.

This study used a mixed-mode design, which means that participants could complete the study via a paper version or an electronic version available on Survey Monkey (2007), a widely used software program designed to allow one to create customized online surveys. The mixed-mode design was implemented in order to maximize the validity of the study. According to Dillman, Smyth, and Christian (2009), a mixed-mode design increases the validity of survey research by increasing the likelihood various types of participants can be contacted and will feel comfortable responding. This study also utilized other strategies to increase the validity of the study (see Appendix A for further information about strategies designed to increase the validity of this study). In addition, in order to estimate response rates and to investigate the reliability of the legal philosophy, role orientations, and FFN scales, a pilot version of this study was conducted. The pilot study resulted in an 11% response rate and the data from these participants were included in the current study. Due to the feedback from, and results of, the pilot study, the wording of the vignettes and scales was modified for the current study (see Appendix B for further information about the pilot study).

Research Questions and Hypotheses

This study sought to answer four research questions regarding risk communication and four research questions regarding judicial attributes. Nine hypotheses were proposed regarding these research questions. Of supplemental research interest was whether these hypotheses were applicable in both low risk and high risk scenarios. The eight research questions and nine hypotheses are delineated below (see Figure 1 for a conceptual model).

- 1) Do judges find some risk messages more probative in value than other risk messages?

Hypothesis 1: Judges will find risk management messages significantly more probative in value than risk prediction or description messages.

Hypothesis 2: Judges will find categorical risk predictions significantly more probative in value than risk prediction messages that use frequency or probabilistic terms.

- 2) Do risk messages impact the restrictiveness of the judicial ruling?

Hypothesis 3: Judges who receive frequency risk prediction messages will be significantly more likely to make a restrictive ruling than judges who receive probability or categorical risk prediction messages.

- 3) Does risk level impact the probative value of risk messages?

Hypothesis 4: Judges who receive the high risk vignettes will be more likely to rate all risk messages as more highly probative in value than will judges who receive the low risk vignettes.

- 4) Do risk messages impact judicial rulings by affecting a judge's determination of dangerousness?

Hypothesis 5: Dangerousness determination will significantly mediate the relation between risk prediction messages (i.e., frequency, categorical and probability messages) and judicial ruling. That is, risk prediction formats will predict dangerousness determinations and, in turn, dangerousness determinations will predict judicial rulings.

- 5) Does FFN impact the restrictiveness of judicial rulings?

Hypothesis 6: The construct FFN will significantly predict judicial ruling. Specifically, as the FFN score increases, the restrictiveness of judicial rulings will also increase.

- 6) Does FFN impact the extent to which risk messages influence judges' commitment decisions?

Hypothesis 7: FFN will moderate the relation between type of risk message and judicial rulings. Specifically, as the FFN score increases, the ability of risk messages to predict ruling will also increase.

7) Does legal philosophy impact the extent to which judges base their judicial ruling on their dangerousness determinations?

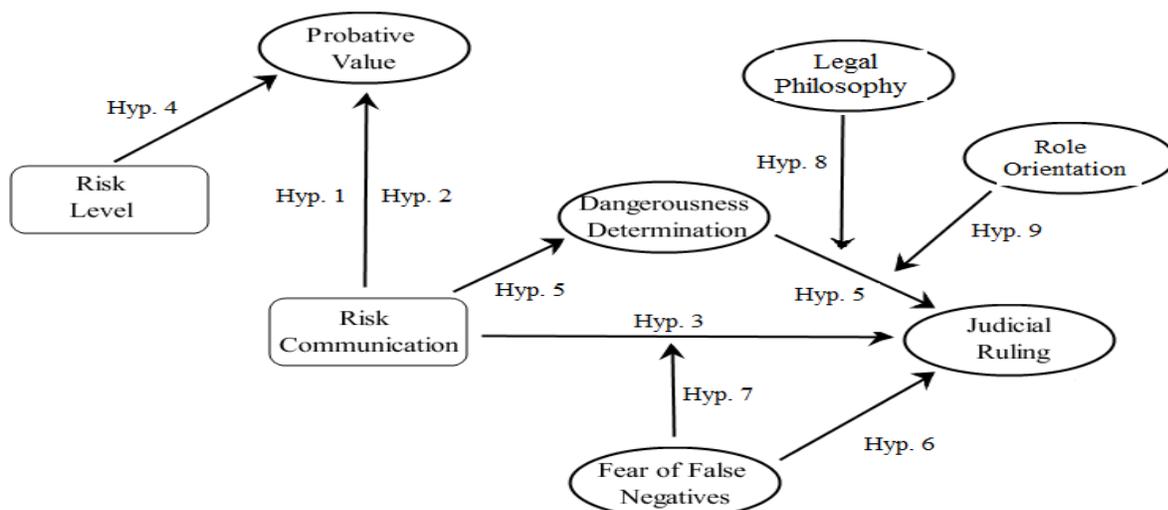
Hypothesis 8: The construct legal philosophy will moderate the relation between dangerousness determinations and judicial rulings. Specifically, as the legal philosophy score increases (i.e., a more police power stance), the ability for the dangerousness determination to predict judicial ruling will also increase.

8) Does role orientation impact the extent to which judges base their judicial ruling on their dangerousness determinations?

Hypothesis 9: The construct role orientation will moderate the relation between dangerousness determinations and judicial rulings. Specifically, as the score on the role orientation scale increases (i.e., more adherence to precedents and statutes) the ability of the dangerousness determination to predict judicial ruling will increase.

Figure 1

Conceptual Model Underlying Hypotheses



The hypotheses of this study were analyzed through analyses of variance (ANOVA) and regression analyses. In situations where the assumptions of an ANOVA were violated, then the nonparametric Kruskal Wallis Test was used. Bonferroni posthoc tests were conducted when an ANOVA resulted in a significant main effect. Ordinal regression analyses were used because the assumptions of linear regression analyses were violated, and subsequent transformations to the data were unable to correct for the violations in normality. In situations where the assumptions of an ordinal regression analysis were violated, a multinomial regression analysis was used instead. It should be noted that risk prediction format and risk models were categorical variables, while FFN, role orientation, and legal philosophy were continuous variables. Judicial ruling, probative value, and dangerousness determination were ordinal variables. Further information about how these constructs were measured is provided in the Measures section.

Participants

Participants of this study were judges who conduct involuntary civil commitments throughout the United States. According to the National Center for State Courts (NCSC, 2009), approximately 12,821 judges met the inclusion criterion for participation in this study. However, information obtained from the pilot study indicated that not all of these judges presided over civil commitment proceedings. Specifically, a judge may belong to a court with jurisdiction over civil commitment, but he or she may be assigned to a department or division that does not oversee civil commitments. In order to maximize the likelihood of obtaining judges who preside over civil commitment proceedings, an extensive review of each state's court structure was conducted and judges who did not oversee civil commitments were excluded from the sample. Information about the court structure was obtained from *The American Bench* (2010) and publically accessible websites (i.e., state or county websites that describe the duties of the judges).

Although demographic information about civil commitment judges was not available, the most recent demographic information about judges, in general, was gathered in 2004 (i.e., Williamson, 2004). According to Williamson (2004), the average judge was 55.1 years old, male (78.5%), and had been on the bench 9.9 years; no information about race was obtained.

In the end, a list of 9,323 judges was generated. This list was thoroughly reviewed² in order to ensure no potential participant was on the list more than once. Removal of duplicates was important, because it decreases coverage error caused by some participants having an increased likelihood of being selected for the study (Dillman et al., 2009). To ensure regional representation, judges in the sample were stratified by U.S. Census Bureau (2009) regions. A breakdown of this list indicated each region had the following number of potential participants: 1,422 in the West; 2,398 in the Midwest; 1,781 in the Northeast; and 3,728 in the South. In order to determine sample size, a power analysis was conducted using a medium effect size, a significance level of $p < .05$, and a power of .80. The results of this power analysis indicated that a sample size of 260 participants (26 per condition) was required to achieve sufficient power.

Participant Recruitment

Potential participants were recruited via email or the United States Postal Service (USPS). Email addresses and work mailing addresses for the judges were obtained through two methods: (a) internet searches via publically accessible websites and (b) searches of directories located at the University of Alabama Law Library. Potential participants were randomly selected using an Excel random number generator method. All judges with email addresses were sent an electronic request for participation letter (see Appendix C); if no email address was available, the

² The list was reviewed by using the “conditional formatting” option of Microsoft Excel. This process resulted in all cells that contained duplicate values being highlighted. The duplicate values were examined and internet searches were conducted in order to ensure each value referred to one judge. Duplicate values that referred to one individual were corrected by deleting one of the entries.

same letter was sent via the USPS (see Appendix D). It should be noted that only 30% (2,855) of the 9,323 eligible judges had email addresses. Thus, the majority of subject recruitment occurred via the USPS. Based on the 11% response rate obtained in the pilot study and the required sample size of 260, 2400 eligible participants were contacted. Two weeks after the first contact email/letter was sent, an additional email or postcard was sent to remind them about the study (see Appendices E and F, respectively). In order to increase response rate and decrease possible nonresponse error (i.e., when participants who respond are different from those who do not respond), participants were given the option to complete the survey electronically or via paper-and-pencil format.

Although inferences about specific regions were not proposed, in order to decrease sampling error the sample was stratified by region (Garson, 2009). The goal was to obtain 64 judges from the West, 65 from the Midwest, 64 from the Northeast, and 65 from the South. The final sample included 131 Western judges, 101 Midwestern judges, 67 Northeastern judges, and 121 Southern judges.

Stratification was accomplished by breaking the recruitment phase into two rounds. For the first round, 1,500 potential participants were randomly selected from the overall pool of 9,323 participants with no attempt to control for region. Specifically, 281 judges from the West, 338 from the Midwest, 303 from the Northeast, and 578 from the South were contacted. Within one month, the required sample size of 65 judges from the Midwest and 65 judges from the South had been obtained, but only 50 participants from the West and 23 from the Northeast had responded (14 were still needed from the West and 41 from the Northeast).

Round two began one month after initial recruitment and consisted of 700 participants being contacted. During this round, the majority of potential participants (i.e., 521 judges) were

selected from the pool of Northeastern judges and 179 were selected from the West. The majority of the potential participants were selected from the Northeastern region, because this region required 75% more participants (41 participants) than the Western region (14 participants). In sum, 200 judges were recruited for the pilot study and 2,200 for rounds one and two of the current study.

Participant Response

Of the 2,400 potential participants identified, contact was made with 2,247. Specifically, 47 letters and 106 emails were returned as undeliverable. Of the 2,247 participants successfully contacted, 43 reported they did not have jurisdiction over involuntary civil commitments, 1,801 did not respond, and 403 chose to participate, resulting in a response rate of 18.28% (403/2204). The lowest response rate was from participants in the pilot study (11%). In the current study, judges who were contacted via USPS had a response rate of 19.77%, while those contacted electronically had a response rate of 16.78%.

The sample size of 403 allowed for a precision of 4.75% when making inferences to the true population. Precision refers to the margin of error and not the confidence level. For example, if 60% of this sample indicated they were females, one could state with 95% confidence that 55.25% to 64.75% of the entire judicial population would be female. According to Salant and Dillman (1994), precision needs to be under 10% to be considered acceptable.

The proposed regional representation, of at least 64 or 65 judges from each region, was achieved. The majority of judges were Southern (121), followed by Western (113) and Midwestern (101), followed by Northeastern (67). Without factoring in whether the potential participant was successfully contacted, the response rate was 18.39% for Southern judges

(121/658), 23% for Western judges (113/490), 28.77% for Midwestern judges (101/351), and 7.75% for Northeastern judges (67 participated /864 contacted). .

Procedure

As previously described, participants were recruited via email or USPS (see Appendices C and D for recruitment email and letter, respectively). Potential participants were informed that the survey would take less than 10 minutes to complete and that their responses would be anonymous.³ Using a randomization formula in Excel (i.e., =RANDBETWEEN[1, 10]), each participant was randomly selected to receive one of ten risk vignettes in which risk message and risk level were systematically varied. Potential participants had the option of completing the study electronically or via a paper-and-pencil version.

If a potential participant chose to participate, he or she completed a questionnaire designed to gather demographic information (see Appendix G). Following the demographic questionnaire, participants read the risk vignette applicable to their condition (see Appendix H). After reading the risk vignette, participants were asked to determine whether they believed the dangerousness criterion had been met, to determine what ruling would be appropriate, and to rate the probative value of the vignette (see Appendix I). Finally, all participants completed a questionnaire designed to assess legal philosophy, role orientation, and FFN (see Appendix J). In addition, some of the participants (excluding those in conditions 3, 4, 8, and 9) were asked five questions designed to provide initial information for a subsequent (and not directly related) study (see Appendix K). These additional questions are not believed to have influenced the responses

³ This statement is supported by the results of four graduate students who completed a longer version of the survey during a manipulation check. Of the four graduate students who participated in the manipulation check, two completed the survey in 6 minutes, one in 8 minutes, and one in 12 minutes. Information regarding the length of time to complete the survey was not obtained from participants in the pilot study.

to the study proper and did not significantly increase the amount of time needed to complete the study (i.e., approximate time to complete this section was one minute).

Measures

All measures were created for use in this study. Because response rates for surveys decrease as the time required to complete surveys increase (Dillman et al., 2009), efforts were directed toward minimizing the number of items needed to adequately measure each construct. In order to decrease the likelihood of measurement error, the following four strategies were implemented. First, the measures contain items or statements that have appeared in the relevant literature. Second, the wording of the items and response options follow guidelines set forth by Dillman and colleagues (2009). Third, a manipulation check with a sample of 17 graduate students and 6 doctoral-level psychologists supported the construct validity of the risk vignettes (see Appendix L for results of manipulation check). Fourth, this study was pilot tested with a sample of 22 judges in order to ensure that the wording of items and instructions were as clear and concise as possible (see Appendix B for results of pilot study).

Demographic questionnaire (see Appendix G). The demographic questionnaire consisted of items related to background variables (i.e., age, gender, ethnic identity, and education), experience as a judge (e.g., years as a judge; number of civil commitments), and number of known prior false negative decisions in civil commitment proceedings.

Risk vignettes (see Appendix H). The vignettes were modeled after risk vignettes used or described in the risk communication literature (Heilbrun et al., 2004; Kwartner et al., 2006; Monahan & Steadman, 1996). This study contained ten separate risk vignettes, one for each of the ten conditions. Within the risk vignettes, risk level and risk message were systematically varied in a 5 x 2 between-groups design. Specifically, there were five types of risk messages and

two risk levels (see Table 2 on p. 25).

Before reviewing the risk vignette, participants received an introductory paragraph that explained the context of the evaluation and informed participants that, with the exception of the dangerousness criterion, all the criteria for civil commitment had been established. Specifically, participants were informed that the respondent had a bonafide mental illness and lacked the capacity to make treatment decisions. Need for treatment was not included in the introduction because feedback from the pilot study indicated this information was unnecessary and at times confusing. In addition, the introduction did not include a reference to danger to self. Although danger to self and danger to others are two distinct concepts, they usually comprise the singular criterion of “dangerousness.” The rationale was that if danger to self was established in the introduction, then judges would have no reason to consider or make a determination on danger to others. In addition, a definitive statement precluding danger to self could have also confounded the results of this study by restricting the judges’ options. For these reasons, danger to self was purposely vague in the introduction to the risk vignettes.

Within the risk vignettes, risk level was manipulated by varying the number of risk factors. Risk vignettes used in previous studies have included zero risk factors as their low risk scenario (Heilbrun et al., 2004; Kwartner et al., 2006; Monahan & Steadman, 1996); however, those risk vignettes did not include any protective factors. Based on the advice of Rogers (2000), protective factors were included in the vignettes used in the current study. As previously noted, Rogers (2000) concluded that a fair and balanced risk assessment requires consideration of protective factors. In the current study, low risk was depicted using one risk factor (i.e., medication noncompliance) and one protective factor (i.e., familial support).

With regard to the high risk vignettes, Kwartner and colleagues (2006) used the presence

of four risk factors (i.e., age less than 25, lengthy criminal record, history of violence toward others, and history of substance use), while Heilbrun and colleagues (2004) and Monahan and Steadman (1996, p. 936) used three risk factors (i.e., substance abuse, medication noncompliance and a history of violence; substance abuse, history of assault, and vague recent threats). For the current study, high risk was depicted by the presence of four risk factors (i.e., age less than 25, medication noncompliance, substance use, and history of violent behavior toward family members and strangers) and one protective factor (i.e., familial support).

In order to establish that the risk vignettes represented the appropriate risk level, a manipulation check was conducted. Seventeen graduate students and six doctoral-level psychologists rated the dangerousness of the respondent using a seven-point Likert scale (1 = *disagree strongly* that the respondent is dangerous to 7 = *strongly agree* that the respondent is dangerous). Mean dangerous rating for the low risk vignette was 3.23 (Mode = 3, *Mdn* = 3, *SD* = 1.59), while the mean for the high risk vignette was 5.54 (Mode = 5 or 6, *Mdn* = 5.5, *SD* = 1.07). The difference between the two groups was statistically significant, $t(21) = 3.70, p = .001$. The results of this manipulation check supported the construct validity of the risk levels.

In addition to risk level, the type of risk communication message was systematically varied in this study. Specifically, there were five types of risk communication messages: (a) risk description model; (b) risk prediction model [categorical format]; (c) risk prediction model [probabilistic format]; (d) risk prediction model [frequency format]; and (e) risk management model. All of the vignettes described the risk and protective factors appropriate to the risk level. Conclusive statements (i.e., prediction of risk or recommendations for intervention strategies) were varied for all groups except those that utilized risk description messages.

In contrast to the risk description or risk management messages, risk prediction messages included a statement indicating the likelihood that the respondent would commit a violent act in the near future. These prediction statements were expressed in one of three ways: (1) categorically, (2) probabilistically, or (3) using frequency terms. The wording of the prediction statements was based on exemplars from the risk communication literature (i.e., Conroy & Murrie, 2007; Heilbrun et al., 2004; Kwartner et al., 2006). The categorical terms used were “low risk” or “high risk.” In order to ensure consistency between the conditions, the probabilistic estimates used in this study were equivalent to the categorical terms. The probabilistic estimates were based on the only study that has obtained the probabilistic equivalent of categorical terms (McNiel and Binder, 1998). McNiel and Binder (1998) found that a 29% likelihood of violence corresponded to low risk and a 76% likelihood corresponded to high risk. The frequency statements were simply the conversion of the probabilities to numerically equivalent frequencies (e.g., 29% equates to 29 out of 100).

The risk management messages consist of recommendations for intervention strategies that may ameliorate dynamic risk factor(s). The intervention strategies used in the current study addressed the dynamic risk factors of medication noncompliance and substance abuse. The other two risk factors (i.e., young age and history of violence) were static and thus, could not be modified. Conroy and Murrie (2007) provided an example of a risk management plan, which included appropriate interventions for medication noncompliance and substance abuse. The authors recommended that medication noncompliance be addressed through “long-acting injectable medication” (p. 143) and substance abuse via random drug screening, attendance at Alcoholics Anonymous (AA), family support, and job selection. To clarify, the low risk vignettes used in this study included the recommendation for “long-acting injectable

medication,” because the low risk vignettes contain the sole risk factor of medication noncompliance. The high risk vignettes contained all of the aforementioned recommendations because these vignettes contained two dynamic risk factors (i.e., medication noncompliance and substance abuse).

Follow-up questions to risk vignettes (see Appendix I). After reading the vignette, participants were asked to answer three questions. Specifically, participants were asked to: (1) determine whether the “danger to others” criterion had been met, (2) determine what ruling was appropriate, and (3) rate the probative value of the risk vignette. The determination of dangerousness was assessed via a seven-point Likert scale (1 = *Strongly disagree* to 7 = *Strongly agree*) that relates to the question: “To what extent do you agree or disagree that the ‘danger to others’ criterion has been sufficiently met?” With regard to judicial ruling, participants used a five-point ordinal scale to indicate which outcome was appropriate (e.g., 1 = *The respondent should not be committed*; 3 = *The respondent should be committed to an outpatient facility*; 5 = *The respondent should be committed to an inpatient facility*). Lastly, participants rated the perceived probative value of the risk vignette using a five-point response scale (1 = *Not at all valuable* to 5 = *Extremely valuable*).

In order to decrease possible measurement error, the wording of these questions and responses followed recommendations set forth by Dillman and colleagues (2010). For example, all questions stated both the negative (e.g., disagree) and positive (e.g., agree) sides in the question stem. According to Dillman and colleagues (2010), this method balances the question so that neither response is given priority. In addition, answer categories included all reasonably possible answers and were mutually exclusive. Finally, the scale responses and answer categories

did not exceed seven options because research has indicated that participants have difficulty processing more than seven options (Dillman et al., 2010).

Judicial attribute questionnaire (see Appendix J). The last measure the participants received was the judicial attribute questionnaire. The purpose of this questionnaire was to assess the constructs of legal philosophy, role orientation, and FFN. This questionnaire consisted of a total of 14 items. Participants were asked to rate their agreement or disagreement of each item through a seven-point Likert scale (1 = *Strongly disagree* to 7 = *Strongly agree*). Further details about each construct, within this questionnaire, is provided below.

Role orientation (see Appendix M items 1-4). The construct role orientation was assessed through a four-item scale. The items used in this scale were based on the items used in the Gibson (1981) scale, but the wording was modified to make the items more applicable to civil commitments. As previously mentioned, participants were asked to respond to each item via a seven-point Likert scale, which was in contrast to the four-point Likert scale (1 = *agree strongly* to 4 = *disagree strongly*) used in the Gibson (1981) scale. With regard to the validity of this scale, results from the pilot study indicated this scale demonstrated adequate internal consistency ($\alpha = .75$) and responses were normally distributed. The possible scores for this scale ranged between 4 and 28. Higher scores indicated that a judge believes that following the letter of the law is important, whereas lower scores indicated that a judge is more likely to believe in the use of discretion.

Legal philosophy (see Appendix M items 5-7). The construct of legal philosophy was assessed through three statements that described the purpose of civil commitment and the pros/cons associated with commitment decisions. Results from the pilot study indicated that this scale demonstrated adequate internal consistency ($\alpha = .7$) and participants' distribution of scores

were normal. The possible scores for this scale ranged between 3 and 21. Higher scores indicated that a judge has a police power stance, whereas lower scores indicated that a judge has a *parens patriae* stance.

Fear of false negatives (FFN; see Appendix M items 8-14). The construct FFN was assessed through seven statements regarding the possible consequences of judicial error. These items were based on statements made in two court cases (i.e., *Addington v. Texas*, 1979; *Goetz v. Crosson*, 1992) and in theoretical discussions of false negatives (i.e., Durham, 1996; Mulvey & Lidz, 1985; Wexler, 1983, 1992; Winick, 2005). Results from the pilot study indicated this scale demonstrated acceptable internal consistency ($\alpha = .69$) and scores were normally distributed. The possible scores for this scale ranged between 7 and 49. Higher scores indicate a judge has a high FFN, whereas lower scores indicate that a judge has a low FFN.

CHAPTER 3

RESULTS

Demographics

Participants of this study came from 45⁴ states and the District of Columbia. As can be seen in Table 3, the average age was 56.79 years old ($SD = 8.51$). Judges reported an average of 12.55 years ($SD = 8.27$) on the bench, with 10.1 years ($SD = 7.66$) with jurisdiction over involuntary civil commitments. Most of the judges were Caucasian (91.9%), male (79.3%), previously employed as attorneys (82.7%), and were functioning in the role of judge at the time they completed the questionnaire (97.2%). The majority of judges had obtained a juris doctor (88.2%) and just over half of them had been elected to the bench (55.8%). In total, the judges had presided over an average estimated total of 240 ($SD = 556.3$)⁵ civil commitment cases, with a mode of 100 cases. The range was very large and went from 1 to 35,000 cases; 8.4% of the judges had signed off on civil commitment cases, but had not personally presided over any of these cases. The majority of judges (29.7%) had personally presided over 100 to 1000 cases and only 10% had presided over more than 1000 cases. The groups did not differ with regard to any demographic variables or the region of the judge, which indicates that participants were adequately randomized into the conditions. (See Appendix N for data regarding the demographic variables broken down by region).

⁴ Delaware, Hawaii, New Hampshire, New Jersey, and Rhode Island were not represented in the sample.

⁵ The presented mean and SD are based on a sample that excluded outliers (i.e., five judges who presided over a total of 7,000 cases were removed). With the inclusion of the outliers, the original mean was 503.97 ($SD = 2,645$).

Table 3
Demographic Variables

Categorical Variables	Frequency	Percentage		
Gender				
Male	315	79.3%		
Female	82	20.7%		
Ethnic Identity				
Caucasian	363	91.9%		
African American	17	4.3%		
Latino/Latina	10	2.5%		
Asian	2	0.5%		
Other	3	0.8%		
Education Level				
High School	2	0.5%		
Some College	12	3.0%		
Bachelor Degree	26	6.5%		
Graduate Degree	7	1.8%		
Juris Doctor (JD)	351	88.2%		
Work Status				
Active Judge	387	97.2%		
Retired Judge	11	2.8%		
Type of Judge				
Magistrate	71	17.8%		
Probate	53	13.3%		
Superior	48	12%		
District	109	27%		
Circuit	69	17.3%		
Court of Common Pleas	15	3.7%		
Other	34	8.6%		
Experience with a known false negative decision				
Yes	31	8.1%		
No	353	91.9%		
Election Status				
Elected	222	55.8%		
Appointed	132	32.8%		
Appointed and then elected	44	10.9%		
Prior experience as attorney				
Yes	330	82.7%		
No	69	17.3%		
Continuous Variables [range]				
Age [24-76]	<i>M</i> 56.79	<i>SD</i> 8.51	Mode 56	<i>Mdn</i> 58
Work Experience				
Years as judge [1-37]	12.55	8.27	4	11
Years as civil commitment judge [0-35]	10.1	7.67	4	9
Number of civil commitment cases [0-35000] ^a	503.97	2645	100	50

^a The presented mean and *SD* are based on entire sample, which includes five outliers (i.e., five judges who presided over 7,000 cases). If the outliers are removed, the mean number of civil commitment cases is 240 (*SD* = 556.3). The median and mode remain the same.

For the most part, judges who responded after initial contact were not significantly different from judges who responded after follow-up contact, with regard to the demographic variables; the exceptions were ethnic identity and gender. Specifically, female ($\chi^2[1, N = 397] = 4.68, p = .03$) and African-American, Asian, Latino, and 'Other' judges ($\chi^2 [4, N = 395] = 14.19, p = .007$) were more likely to respond during the second contact phase than during the first. However, these differences did not reach statistical significance when accounting for the Bonferroni correction p – value of .0028 (.05/18). Response phase did not impact any of the outcome or manipulation variables.

Results from a chi-square analysis indicated that the majority of judges (79.2%) who responded after the first contact completed the paper version of the study instead of the electronic version, $\chi^2(1, N = 403) = 16.04, p < .001$. With regard to follow-up contact, only 19.7% of the judges who received the USPS letter completed the paper version, while 39.4% of judges who received the electronic recruitment email completed the electronic version of the study. Thus, electronic follow-up resulted in an increased percentage of electronic participation, while USPS follow-up did not substantially increase paper-version participation. Judges who responded via the electronic version of the study did not differ from those who responded to the paper version of the study with regard to the demographic variables, conditions assigned, or the outcome variables.

Hypotheses

Risk communication.

Hypothesis one: Based on the results of this study, the hypothesis that judges would rate risk management messages as significantly more probative than other messages (i.e., description or prediction) was rejected. A one-way between subjects ANOVA was used to analyze this hypothesis and the assumptions of normality and equal variance were not violated. Results from the one-way between subjects ANOVA indicated that judges did not rate any one of the three risk models as more probative, $F(2, 374) = 1.5, p = .23$; however, prediction messages ($M = 3.26, SD = 1.57$) and management messages ($M = 3.23, SD = 1.11$) obtained slightly higher probative scores than did description messages ($M = 3.0, SD = 1.08$). As previously mentioned, probative value was determined by responses on a five-point ordinal scale in which higher scores indicated that a judge perceived the risk vignette to have higher probative value (1 = *Not at all valuable* to 5 = *Extremely valuable*); three reflects “somewhat valuable” (see Table 4 for the frequency with which judges endorsed each response). Further analyses indicated risk level did not moderate the relation between risk model and probative value, $F(2, 376) = 0.18, p = .84$ (see Figure 2 for mean probative value varied by risk model and risk level). Specifically, risk message did not impact probative value in the high risk ($F[2, 186] = 1.85, p = .16$), or low risk scenarios ($F[2, 185] = 0.84, p = .43$).

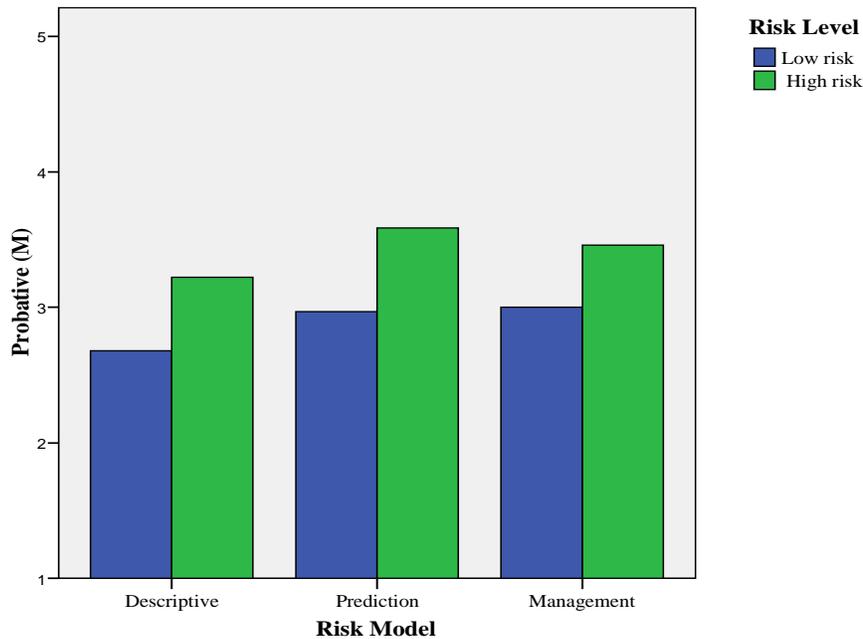
Table 4

Frequency of Probative Rating Endorsement by Risk Model

Probative Value	Risk Description ($N = 61$)	Risk Prediction ($N = 242$)	Risk Management ($N = 74$)
1. Not at all valuable	8.2%	6.6%	8.1%
2. Slightly valuable	24.6%	14.5%	13.5%
3. Somewhat valuable	34.4%	36.4%	39.2%
4. Very valuable	24.6%	31.0%	25.7%
5. Extremely valuable	8.2%	11.6%	13.5%

Figure 2

Mean of Probative Value varied by Risk Model and Risk Level



Hypothesis two. Based on the results of this study, the following hypothesis was supported: Judges rated risk prediction messages that used categorical terms as significantly more probative than other risk prediction messages. A one-way between subjects ANOVA showed that judges' perception of probative value differed, depending on the way the risk prediction was communicated, $F(2, 239) = 4.55, p = .01^6$ (see Table 5 for the frequency with which judges endorsed each response). A Bonferroni posthoc analysis indicated that judges find risk prediction messages expressed via categorical terms ($M = 3.54, SD = 0.92$) to be significantly more probative in value than risk prediction messages that use frequency ($M = 3.04, SD = 1.13$) or probabilistic ($M = 3.15, SD = 1.07$) terms (see Table 6 for the statistical information of each Bonferroni comparison). A between subjects multivariate analysis of variance (MANOVA) indicated that risk level did not significantly moderate the relation

⁶ This hypothesis was also supported when analyzed via a between subjects multivariate analysis of variance (MANOVA) that included the dependent variables of probative value and judicial ruling, along with the independent variables of risk level and risk prediction format, $F(2, 236) = 5.03, p = .007, \text{Partial } \eta^2 = .04$. This finding decreased the likelihood the acceptance of this hypothesis was an example of a Type I error.

between risk prediction format and probative value, $F(2, 242) = 0.299, p = .74$ (see Figure 3 for mean of probative value varied by risk format and risk level).

Table 5

Frequency of Probative Rating Endorsement by Risk Format

Probative Value	Categorical Terms (<i>N</i> = 83)	Frequency Terms (<i>N</i> = 78)	Probabilistic Terms (<i>N</i> = 81)
1. Not at all valuable	3.6%	9.0%	7.4%
2. Slightly valuable	6.0%	19.2%	18.5%
3. Somewhat valuable	34.9%	38.2%	35.8%
4. Very valuable	43.4%	20.5%	28.4%
5. Extremely valuable	12.0%	12.8%	9.9%

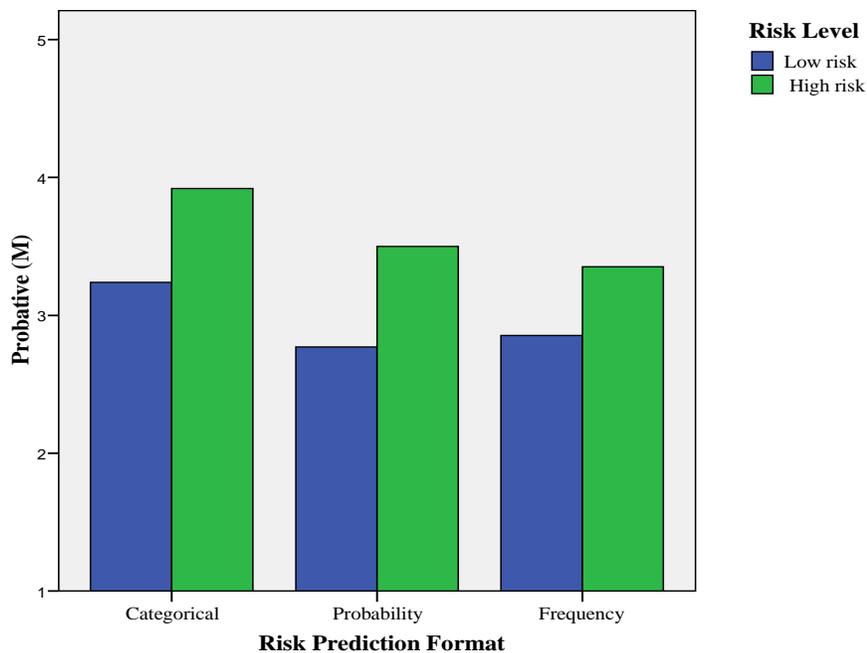
Table 6

Mean Difference, Standard Error, and Significance of Bonferroni Comparisons

Comparison	Mean Difference	<i>SE</i>	<i>p</i>
Categorical vs. Probability	.39	.16	.048
Categorical vs. Frequency	.45	.16	.019
Probability vs. Frequency	.06	.17	1.00

Figure 3

Mean of Probative Value varied by Risk Format and Risk Level



Hypothesis three. Based on the results of this study, the hypothesis that judges who receive risk prediction messages in the form of frequencies would make more restrictive rulings⁷ was rejected. A review of the assumptions of an ANOVA indicated that judicial ruling was not normally distributed ($M = 3.27$, Mode = 5, $Mdn = 4$). A histogram indicated that judicial ruling was negatively skewed, with a standardized skew of -2.82 (skew = -0.46, $SE = 0.16$) and platykurtic, with a standardized kurtosis of -3.35 (kurtosis = -1.09, $SE = 0.33$). A Kruskal Wallis test was used instead of an ANOVA because the assumption of normality was violated. Results from a Kruskal Wallis analysis indicated that risk prediction format did not influence the restrictiveness of a judge's ruling, $\chi^2(2, N = 238) = 1.81, p = .41$. Judges who received categorical risk messages made slightly less restrictive rulings ($M = 3.27, SD = 1.48$) than did judges who received either frequency messages ($M = 3.48, SD = 1.42$) or probability messages ($M = 3.58, SD = 1.39$). Risk level, low versus high, did not moderate this relation ($\chi^2[2, N = 122] = 2.99, p = .22$ and $\chi^2[2, N = 116] = 0.04, p = .98, respectively$). It should be noted that investigations into different conceptualization of prediction format (i.e., categorical messages versus numerical messages) and ruling restrictiveness (i.e., not committed versus inpatient commitment) led to the same conclusion: risk prediction format does not impact ruling restrictiveness (see Table 7 for statistical information for these different conceptualizations).

Table 7

Kruskal Wallis Statistics for Different Conceptualizations of Risk Format and Ruling

Risk Format	Ruling	χ^2	df	p
Categorical vs. Numerical ^a	Five-point Ordinal Scale	1.62	1	.20
Categorical vs. Numerical ^a	Dichotomous (not committed vs. inpatient)	1.87	1	.17
Original three formats (Cat., Prob., and Freq.)	Dichotomous (not committed vs. inpatient)	1.87	2	.39

^a The numerical sample refers to the probabilistic and frequency data collapsed into one sample.

⁷ As previously mentioned, judicial ruling was determined by responses on a five-point ordinal scale that went from least restrictive to most restrictive recommendations (i.e., 1 = No civil commitment to 5 = Inpatient civil commitment).

The methodology of this study allowed for an investigation into the impact of risk prediction format on the accuracy of the commitment decision. Objectively, a low risk vignette was meant to represent a respondent who is not dangerous and does not require civil commitment (i.e., true negative). The high risk vignette was designed to represent a respondent who is dangerous and requires inpatient commitment (i.e., true positive). An investigation into the impact of risk messages on the accuracy of judicial decisions was made possible when the data sample was narrowed down to judges who unequivocally decided the respondent did not require commitment (score = 1; $N = 39$) versus those who recommended inpatient commitment (score = 5; $N = 72$). As can be seen in Table 8, categorical messages resulted in a higher sensitivity, specificity, positive predictive power (PPP), negative predictive power (NPP), and hit rates in judicial decisions, as compared to numerical messages. Receiver Operating Characteristic (ROC) analysis was used to analyze the ability of the risk prediction format to help judges discriminate between respondents who required inpatient commitment from those that did not. The Area Under the Curve (AUC) for judges who received the categorical messages was .9, while the AUC for judges who received the numerical messages was .77. A value of 1 for the AUC represents perfect (100%) accuracy, whereas a value of .50 represents chance alone.

Table 8

Classification Accuracy of Judicial Decisions by Prediction Format

Prediction Format	Sensitivity	Specificity	PPP	NPP	Hit rate	AUC
Categorical	.95	.84	.86	.94	.90	.90
Numerical ^a	.89	.65	.82	.72	.81	.77
Frequency	.91	.69	.84	.82	.83	.80
Probability	.87	.69	.80	.73	.78	.74

Note: Sensitivity refers to the proportion of judges who received the high risk vignettes and correctly decided that inpatient commitment was warranted. Specificity refers to the proportion of judges who received the low risk vignettes and correctly decided that commitment was not warranted. Positive Predictive Power (PPP) refers to the proportion of judges who correctly decided inpatient commitment was warranted. Negative Predictive Power (NPP) refers to the proportion of judges who correctly decided that commitment was not warranted. Hit rate refers to the proportion of judges who correctly decided inpatient commitment was warranted or who correctly decided that commitment was not warranted.

^a Numerical refers to judges who received either the frequency or probability messages.

The hypothesis that frequency messages would result in more restrictive rulings was proposed to answer the following research question: Do risk messages impact the restrictiveness of the judicial ruling? In this study, risk prediction format (i.e., categorical, frequency, and probability) was utilized as the conceptualization of risk messages because there was a literature base to support the hypothesis that frequency messages would result in restrictive rulings (Monahan et al., 2002; Slovic et al., 2000). However, of interest was whether a reconceptualization of risk messages would provide a different answer to the aforementioned research question. Instead of the use of prediction messages, risk model was utilized to address the research question. The risk model contained the following three levels (see Table 2 for a reminder of all conditions): description messages (condition one and six), prediction messages (condition two, three, four, seven, eight, and nine), and management messages (condition five and ten).

Due to a violation of the normal distribution, a Kruskal Wallis analysis was used to investigate the impact of risk model on judicial ruling. Results from the Kruskal Wallis indicated that risk model significantly influenced the restrictiveness of a judge's ruling, $\chi^2(2, N = 373) =$

7.94, $p = .019$. Further analysis indicated that judges who received risk prediction messages ($M = 3.44$, $SD = 1.43$) were more likely to make stricter rulings, than were judges who received risk description messages ($M = 2.97$, $SD = 1.52$; $\chi^2[1, N = 299] = 4.92$, $p = .027$) or risk management messages ($M = 2.99$, $SD = 1.57$; $\chi^2[2, N = 312] = 4.78$, $p = .029$). This pattern of results was found in both the low risk ($\chi^2[2, N = 184] = 10.93$, $p = .004$) and high risk scenarios ($\chi^2[2, N = 189] = 6.64$, $p = .036$).

Hypothesis four. Based on the results of this study, the hypothesis that judges would rate high risk messages as significantly more probative than low risk messages was supported. Results from a one-way between subjects ANOVA indicated that the risk level of the vignette had a significant effect on the probative value assigned by judges, $F(1, 375) = 27.14$, $p < .001$ ⁸. Specifically, judges rated the high risk messages ($M = 3.49$, $SD = 1.00$) as more probative in value than the low risk messages ($M = 2.94$, $SD = 1.07$).

Although judges rated all the risk messages to be more probative in high risk scenarios than in low risk scenarios, the strength of this relation differed depending on the risk message the judge received (see Figure 4 for illustration of these relations). This relation was strongest for judges who received categorical prediction messages and weakest for those who received risk management messages (see Table 9 for ANOVA results). The majority of judges (41.3%) who received the categorical message with the term “high risk” rated these messages as “very valuable” ($M = 3.92$, Mode = 4, $Mdn = 4$, $SD = 0.68$), while the majority (54.1%) of those who received the message with the term “low risk” rated the messages as “somewhat valuable” ($M = 3.24$, Mode = 3, $Mdn = 3$, $SD = 0.97$).

⁸ This hypothesis was also supported when analyzed via a between subjects MANOVA that included the dependent variables of probative value and judicial ruling, and the independent variables of risk level, risk prediction format, and an interaction between the two, $F(1, 236) = 22.57$, $p < .001$, Partial $\eta^2 = .09$; and in a MANOVA that included the dependent variables of probative value and judicial ruling and the independent variables of risk level, risk model, and an interaction between the two, $F(1, 369) = 17.2$, $p < .001$, Partial $\eta^2 = .05$

Figure 4

Impact of Risk Level on Probative Value varied by Risk Message

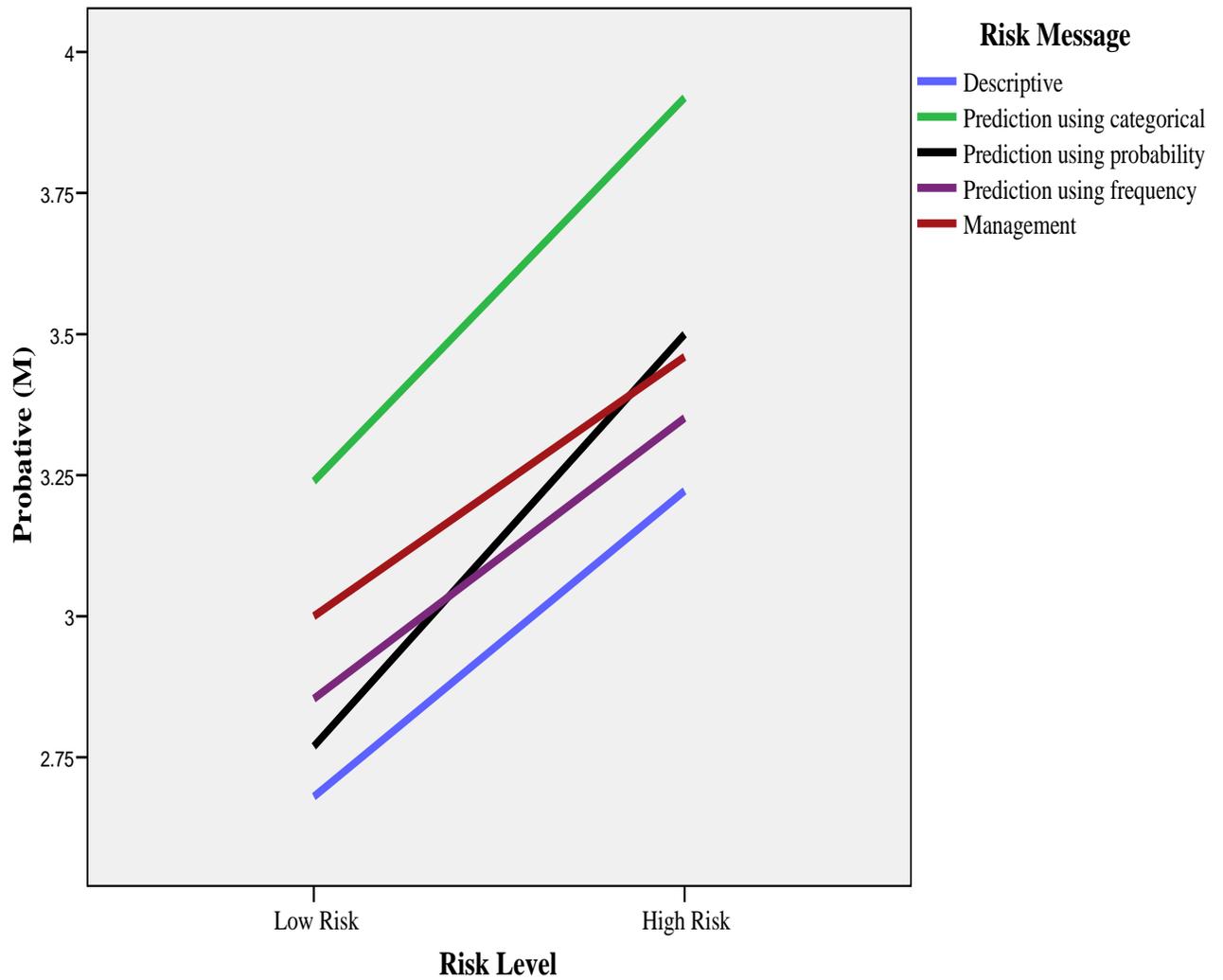


Table 9

Analysis of Variance for Risk Level predicting Probative Value by Risk Message

Risk Message	Low Risk		High Risk		<i>df</i>	<i>F</i>	Partial η^2	Adjusted r^2	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>					
Description Message	2.68	0.9	3.22	1.15	59	3.90	.06	.05	.05
Prediction using categorical terms	3.24	0.97	3.92	0.68	81	13.00	.14	.13	.001
Prediction using probabilistic terms	2.77	1.11	3.50	0.92	79	10.48	.11	.11	.002
Prediction using frequency terms	2.85	1.04	3.35	1.18	76	3.91	.05	.04	.05
Management	3.00	1.23	3.46	0.93	72	3.30	.04	.03	.07

Hypothesis five. It was hypothesized that dangerousness determination would mediate the relation between risk prediction format (i.e., frequency, categorical and probability messages) and judicial ruling. To clarify, dangerousness was determined by the extent to which the judge agreed or disagreed that the dangerousness criterion has been met (i.e., via a seven-point Likert scale); with higher scores associated with higher agreement that the respondent was dangerousness. Results from a Sobel Test indicated that this mediation relation was not supported (e.g., $Z = 0.99$, $SE = 0.15$, $p = .32$, 95% CI [-0.14, 0.44] for categorical versus probability messages).⁹

In order for mediation to be established, the following three conditions must be met (Baron & Kenny, 1986). First, the independent variable (i.e., risk prediction format) must predict the dependent variable (i.e., judicial ruling), which indicates there is a relation that may be mediated. As discussed in hypothesis three (see pp. 47 - 50), risk prediction format had no impact on judicial ruling, $\chi^2(2, N = 238) = 1.81$, $p = .41$. The second condition is that the independent variable (i.e., risk prediction format) must predict the mediator variable (i.e., dangerousness determination). Results from an ordinal regression indicated there was no relation between risk prediction format and dangerousness, $\chi^2(2, N = 244) = 1.34$, $p = .51$. Finally, the mediator (i.e., dangerousness determination) should impact the dependent variable (i.e., judicial ruling). The only association that was significant in this mediation analysis was the association between dangerous determination and judicial ruling ($\chi^2[24, N = 371] = 263.64$, $p < .001$, $r^2 = .53$), with more restrictive rulings being associated with higher levels of agreement that the respondent was dangerous.

⁹ Risk Prediction Format is a categorical variable that had to be dummy coded in order to be included in the regression analyses. For this reason, three separate Sobel tests were calculated for this mediation hypothesis (i.e., for categorical vs. probabilistic, for categorical vs. frequency, and for frequency vs. probabilistic). For ease of read, only the results of categorical vs. probabilistic will be presented in the text, but the other two results did not differ in statistical significance (categorical vs. frequency or frequency vs. probabilistic). The mediation analysis was conducted by using a SPSS macro written by Andrew F. Hayes and available at <http://www.afhayes.com/spss-sas-and-mplus-macros-and-code.html>.

Further analyses indicated that risk level may have complicated this mediation analysis. Specifically, results from Kruskal Wallis analyses indicated that risk prediction format influenced dangerousness determination in both the low risk and high risk conditions ($\chi^2[2, N = 126] = 7.36, p = .025$ and $\chi^2[2, N = 118] = 6.69, p = .035, respectively$). However, the relations were opposite in direction. In the low risk conditions, judges who received the categorical messages ($M = 2.37, SD = 1.76$) were least likely to rate the respondent as dangerous, compared to judges who received the probability ($M = 3.28, SD = 2.03$) and frequency messages ($M = 3.29, SD = 1.98$). In the high risk conditions, judges who received the categorical messages ($M = 6.21, SD = 1.38$) were the most likely to rate the respondent as dangerous, compared to judges who received the probability ($M = 5.76, SD = 1.62$) and frequency messages ($M = 5.45, SD = 1.69$). Thus, when the low and high risk conditions were collapsed into one analysis, the relation between risk prediction format and dangerousness became undetectable.

Due to this additional information, the high risk sample was separated from the low risk sample in order to reanalyze the hypothesis that dangerousness determination would mediate the relation between risk prediction format and judicial ruling. For the low risk sample, risk prediction format predicted dangerousness determinations ($\chi^2[2, N = 126] = 7.62, p = .02$), which in turn predicted judicial ruling ($\chi^2[6, N = 183] = 56.66, p < .001$). Although risk prediction format did not significantly impact judicial ruling (see hypothesis three), when dangerousness determination was controlled for, the relation between prediction format and ruling weakened even further (see Table 10 for Sobel Test results). With regard to the high risk sample, risk prediction format also predicted dangerousness determinations ($\chi^2[2, N = 118] = 6.95, p = .03$), which in turn predicted judicial ruling ($\chi^2[6, N = 188] = 75.64, p < .001$). However, when dangerousness determination was controlled for, the nonsignificant relation between prediction format and ruling did not substantially weaken.

Table 10

Dangerousness as a Mediator between Risk Format and Ruling varied by Risk Level

Risk Level	Risk Prediction Format	Z	SE	95% CI LL	95% CI UL	p
Low Risk	Categorical v. Frequency	2.33	0.17	.06	.73	.02
	Categorical v. Probabilistic	1.62	0.12	-.04	.44	.11
	Frequency v. Probabilistic	-0.52	0.12	-.31	.18	.60
	Categorical v. Numerical	2.28	0.12	.04	.5	.02
High Risk	Categorical v. Frequency	-1.86	0.12	-.46	.01	.06
	Categorical v. Probabilistic	-1.43	0.14	-.48	.08	.15
	Frequency v. Probabilistic	0.66	0.16	-.20	.41	.51
	Categorical v. Numerical	-1.93	0.12	-.47	.003	.05

Similar results were obtained when risk message was reconceptualized from risk prediction formats to risk models. For the overall sample, dangerousness did not mediate the relation between risk model and judicial ruling (e.g., $Z = 1.75$, $SE = 0.14$, $p = .08$, 95% CI [-0.03, 0.51] for description versus prediction messages). Although risk model significantly impacted judicial rulings ($\chi^2[2, N = 373] = 7.94$, $p = .019$), risk model had no impact on dangerousness determinations ($\chi^2[2, N = 379] = 5.17$, $p = .075$). However, a significant mediation effect was observed when the high risk sample was analyzed separately from the low risk sample; specifically, in the high risk sample, risk model predicted dangerousness determinations ($\chi^2[2, N = 190] = 9.51$, $p = .009$), which in turn predicted judicial ruling ($\chi^2[6, N = 188] = 75.64$, $p < .001$). In addition, the relation between risk model and judicial ruling became nonsignificant when the mediator of dangerousness determination was controlled (see Table 11 for Sobel Test results).

Table 11

Dangerousness as a Mediator between Risk Model and Ruling varied by Risk Level

Risk Level	Risk Model	Z	SE	95% CI LL	95% CI UL	p
Low Risk	Description v. Prediction	1.62	.13	-.04	.46	.11
	Description v. Management	0.88	.17	-.19	.49	.38
	Prediction v. Management	-0.83	.12	-.34	.14	.41
High Risk	Description v. Prediction	2.8	.13	.11	.62	.005
	Description v. Management	0.61	.21	-.29	.55	.54
	Prediction v. Management	-1.98	.13	-.53	-.003	.048

Judicial attributes.

Hypothesis six. The hypothesis that Fear of False Negatives (FFN) would predict judicial ruling was supported by the data. With regard to the seven-item FFN scale, participants in this study obtained scores between 13 and 48 (total range possible = 7 to 49), with an average score of 28.92 (Mode = 29, *Mdn* = 29, *SD* = 6.28; see Figure 5 for distribution of FFN scores). The internal consistency of this scale ($\alpha = .54$) was inadequate (i.e., $< .7$), but Cronbach alpha coefficients of .5 are reported to be common in scales that have less than ten items (Pallant, 2007).

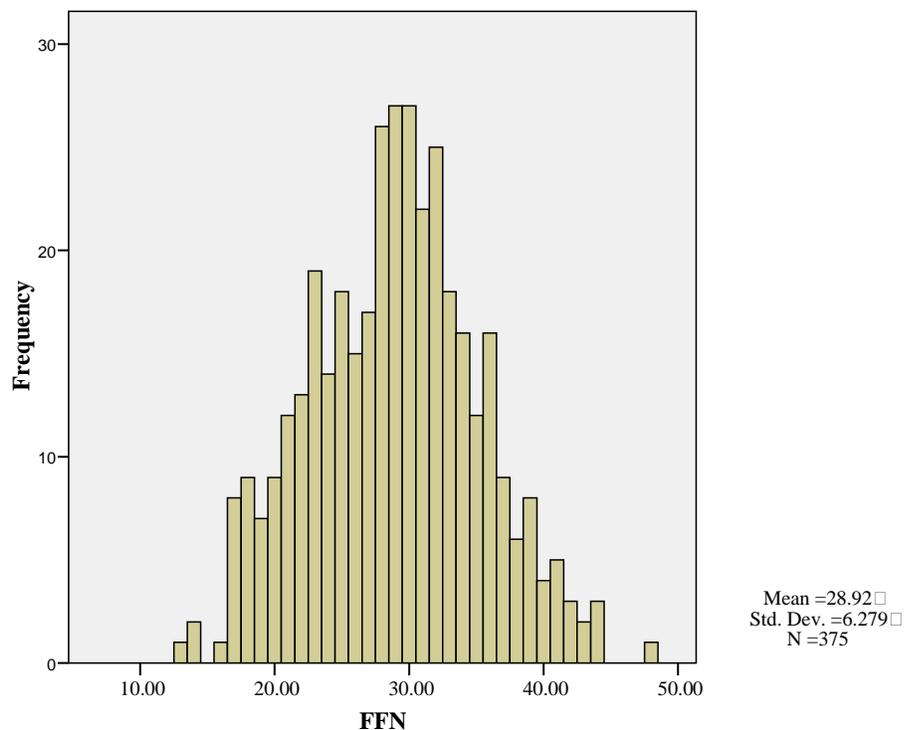
Results from an ordinal regression analysis indicated that, as FFN increased, so did the restrictiveness of the judicial ruling, $\beta = 0.05$, $SE = 0.02$, Wald = 10.68, $df = 1$, $p = .001$, Pseudo $R^2 = .03$.¹⁰ For example, judges who recommended that the respondent not be committed obtained significantly lower FFN scores ($M = 27.11$, $SD = 6.27$) than did judges who recommended that respondent be committed ($M = 30.3$, $SD = 5.99$; $\chi^2[4, N = 364] = 11.63$, $p = .02$). FFN impacted the restrictiveness of judicial ruling in both low risk ($\beta = 0.05$, $SE = 0.02$, Wald = 6.38, $df = 1$, $p = .01$, Pseudo $R^2 = .036$) and high risk scenarios ($\beta = 0.07$, $SE = 0.025$, Wald = 7.16, $df = 1$, $p = .007$,

¹⁰ This hypothesis was also supported when analyzed via an ordinal regression model that included the independent variables of risk level, risk prediction format, dangerousness determination, FFN, legal philosophy, role orientation, and an interaction between legal philosophy and dangerousness determination, $\beta = 0.07$, $SE = 0.02$, Wald = 10.75, $df = 1$, $p = .001$. This finding decreased the likelihood the acceptance of this hypothesis was an example of a Type I error.

Pseudo $R^2 = .044$). In addition, FFN impacted both the restrictiveness of judicial ruling ($\beta = 0.05$, $SE = 0.02$, $Wald = 8.998$, $df = 1$, $p = .003$) and dangerousness determinations ($\beta = .033$, $SE = .015$, $Wald = 4.64$, $df = 1$, $p = .031$), even when the other variables (i.e., risk level, legal philosophy, and role orientation) were controlled for in the model.

Figure 5

Distribution of FFN scores



The methodology of this study allowed for the investigation into the impact of FFN on the accuracy of the commitment decision. As previously mentioned in the discussion of hypothesis three, a low risk vignette was meant to represent a respondent who does not require civil commitment (i.e., true negative), while the high risk vignette was designed to represent a respondent who requires inpatient commitment (i.e., true positive). The data sample was narrowed down to judges who unequivocally decided to not commit the respondent (score = 1; $N = 76$) versus those who

recommended inpatient commitment (score = 5; $N = 99$). In addition, the judges with FFN scores below the mean (28.92) were conceptualized as “low FFN,” while judges with FFN scores above the mean were conceptualized as “high FFN.” As can be seen in Table 12, low FFN judges exhibited higher sensitivity, specificity, positive predictive power (PPP), negative predictive power (NPP), and hit rates in judicial decisions, as compared to high FFN judges.

Table 12
Classification Accuracy of Judicial Decisions by FFN levels

FFN level	Sensitivity	Specificity	PPP	NPP	Hit rate	AUC
Low FFN	.85	.86	.85	.86	.85	.86
High FFN	.83	.27	.69	.47	.64	.78

Note: Sensitivity refers to the proportion of judges who received the high risk vignettes and correctly decided that inpatient commitment was warranted. Specificity refers to the proportion of judges who received the low risk vignettes and correctly decided that commitment was not warranted. Positive Predictive Power (PPP) refers to the proportion of judges who correctly decided inpatient commitment was warranted. Negative Predictive Power (NPP) refers to the proportion of judges who correctly decided that commitment was not warranted. Hit rate refers to the proportion of judges who correctly decided inpatient commitment was warranted or who correctly decided that commitment was not warranted.

Hypothesis seven. The hypothesis that FFN would moderate the relation between risk prediction format and judicial ruling was rejected. In order to investigate this hypothesis, an ordinal regression model was created in which FFN, risk prediction, and an interaction between the two, were independent variables while judicial ruling was the dependent variable. It should be noted that, in order to decrease multicollinearity, FFN was centered in the interaction. Results indicated that FFN does not moderate the relation between risk prediction format and judicial ruling ($\beta = 0.02$, $SE = 0.04$, $Wald = 0.14$, $df = 1$, $p = .71$ for categorical versus frequency and $\beta = 0.06$, $SE = 0.05$, $Wald = 1.47$, $df = 1$, $p = .23$ for probabilistic versus frequency). That is, judges with higher levels of FFN were no more likely to be influenced by risk prediction formats ($\chi^2[4, N = 113] = 4.68$, $p = .32$), than

were judges with lower levels of FFN ($\chi^2[4, N = 123] = 4.63, p = .33$). Risk level did not moderate this relation (see Table 13 for statistical information about these analyses). In addition, actual experience with false negative errors did not moderate this relation, $\chi^2[8, N = 229] = 9.29, p = .32$.

Table 13

Parameter Estimate Statistics for Interactions between Risk Level, FFN, and Risk Format

Three-Way Interaction	β	<i>SE</i>	Wald	<i>df</i>	<i>p</i>
Categorical x FFN x Risk Level	.01	.06	0.02	1	.88
Probabilistic x FFN x Risk Level	.10	.08	1.65	1	.20
Frequency x FFN x Risk Level	.09	.08	1.26	1	.26

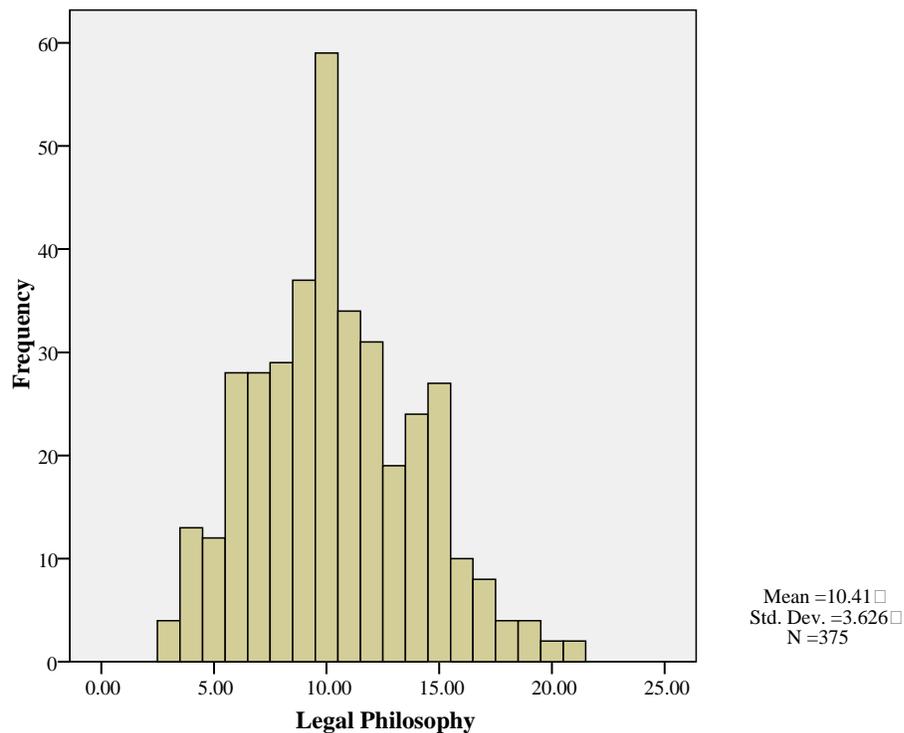
The hypothesis that FFN would moderate the relation between risk messages and judicial ruling was reanalyzed with different conceptualizations of risk message and judicial ruling. There was no significant moderation effect when using risk model instead of risk prediction format ($\beta = -0.98, SE = 0.06, Wald = 3.04, df = 1, p = .08$ for description versus management and $\beta = -0.39, SE = 0.04, Wald = .83, df = 1, p = .36$ for prediction versus management). In addition, there was no significant moderation effect when a dichotomous conceptualization of judicial ruling (i.e., only judges who recommended the respondent not be committed versus those who recommended inpatient commitment) was used instead of the five-point ordinal version of judicial ruling ($F[2, 107] = 0.17, p = .87$ for Risk Format and $F[2, 172] = 0.49, p = .62$ for Risk Model).¹¹ Similarly, FFN did not moderate the relation between dangerousness determinations and the five-point ordinal version of judicial ruling ($\beta = 0.02, SE = 0.07, Wald = .09, df = 1, p = .76$).

¹¹ Two-way between subjects ANOVAs were used to analyze: 1) whether an interaction between risk format and FFN had an impact on a dichotomous conceptualization of judicial ruling; and 2) whether an interaction between risk model and FFN had an impact on a dichotomous conceptualization of judicial ruling. The assumptions of an ANOVA were not violated.

Hypothesis eight. The hypothesis that, as legal philosophy scores increased the relation between dangerous determination and judicial ruling would strengthen, was supported. With regard to the three-item legal philosophy scale, participants in this study obtained scores between 3 and 21 (total range possible = 3 to 21), with the average score of 10.41 (Mode = 10, *Mdn* = 10, *SD* = 3.62; see Figure 6 for distribution of Legal Philosophy scores). Higher scores on this scale indicated that a judge has a police power stance, whereas lower scores indicated that a judge has a *parens patriae* stance. The internal consistency of this scale ($\alpha = .36$) was poor, which lowered the likelihood a significant effect could be detected.

Figure 6

Distribution of Legal Philosophy Scores



To test this hypothesis, an ordinal regression was conducted and this model did not violate the assumption of parallel lines. In order to decrease multicollinearity, legal philosophy was centered in

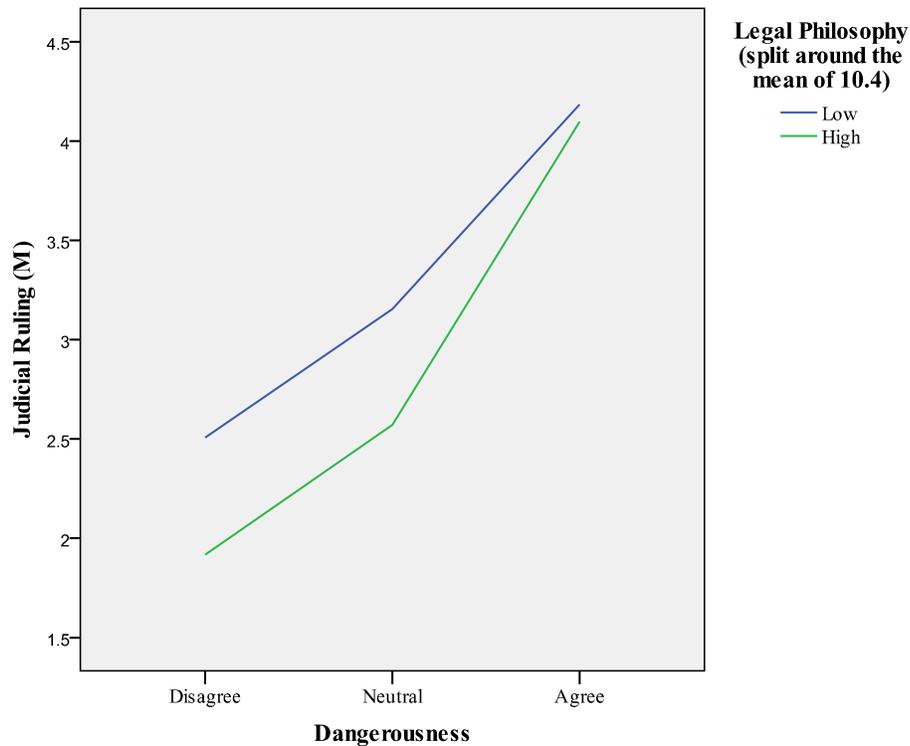
the interaction. The regression model produced a significant interaction effect, which indicated that legal philosophy moderates the relation between dangerousness determination and judicial ruling, $\beta = 0.03$, $SE = 0.01$, Wald = 4.57, $df = 1$, $p = .03$.¹²

As can be seen in Figure 7, judges with *parens patriae* stance (i.e., lower legal philosophy scores) were significantly more likely than were judges with a police power stance to provide a restrictive ruling, even when they opined that the dangerousness criterion had not been established ($\chi^2[1, N = 150] = 8.15$, $p = .004$). Specifically, of the 99 judges who recommended “inpatient commitment,” 65.7% of them had a *parens patriae* philosophy (i.e., below the mean of 10.4 on the legal philosophy scale). Additionally, of the 10 judges who chose inpatient commitment and opined the respondent was not dangerousness, 70% of them had a *parens patriae* stance. Further information about the strength of this moderation effect was obtained when the data sample was divided by legal philosophy score (e.g., above the mean represented police power judges; below the mean represented *parens patriae*) and two ordinal regression analyses were conducted. Specifically, the relation between dangerousness determination and judicial ruling was stronger (Pseudo $R^2 = .51$) for judges with police power stance ($\beta = 0.79$, $SE = 0.09$, Wald = 80.18, $df = 1$, $p < .001$) than was the strength of the relation between dangerousness determination and judicial ruling (Pseudo $R^2 = .41$) for *parens patriae* judges ($\beta = 0.66$, $SE = 0.07$, Wald = 84.8, $df = 1$, $p < .001$). Risk level did not significantly impact this moderation effect.

¹² This hypothesis was also supported when analyzed via an ordinal regression model that included risk level, risk prediction format, dangerousness determination, FFN, legal philosophy, role orientation, and an interaction between legal philosophy and dangerousness determination, $\beta = 0.03$, $SE = 0.02$, Wald = 3.94, $df = 1$, $p = .047$. This finding decreased the likelihood the acceptance of this hypothesis was an example of a Type I error.

Figure 7

Legal Philosophy as a Moderator between Dangerous Determination and Ruling



Note. Judicial ruling was determined by responses on a five-point ordinal scale that went from least restrictive to most restrictive recommendations (i.e., 1 = *The respondent should not be committed* to 5 = *The respondent should be committed to an inpatient facility*). Dangerousness was determined by responses on a seven-point Likert Scale (1 = *Strongly disagree* to 7 = *Strongly agree*) asked participants to rate their disagreement or agreement that the ‘danger to others’ criterion has been sufficiently met?” Only three value points of dangerousness (*disagree* equates to scores of 1-3, *neutral* to a score of 4, and *agree* to scores of 5-7) are presented for visual clarity.

Although the moderating impact of legal philosophy was small, the results of a multinomial regression model (which controlled for risk level and the other attribute variables) demonstrated that legal philosophy was the attribute that most increased the odds that a judge would decide that inpatient commitment was warranted, $\chi^2(4, N = 343) = 21.61, p < .001$. Specifically, the odds of a judge ruling that inpatient commitment was warranted (versus no commitment) increased by a factor of 1.23 for each unit decrease in his or her legal philosophy score (i.e., more likely to hold a *parens patriae* stance), $\beta = 0.21, SE = 0.06, Wald = 14.38, df = 1, p < .001$. In contrast, for each unit increase in FFN, the odds the judge would make an inpatient commitment decision, increased by a factor 0.91

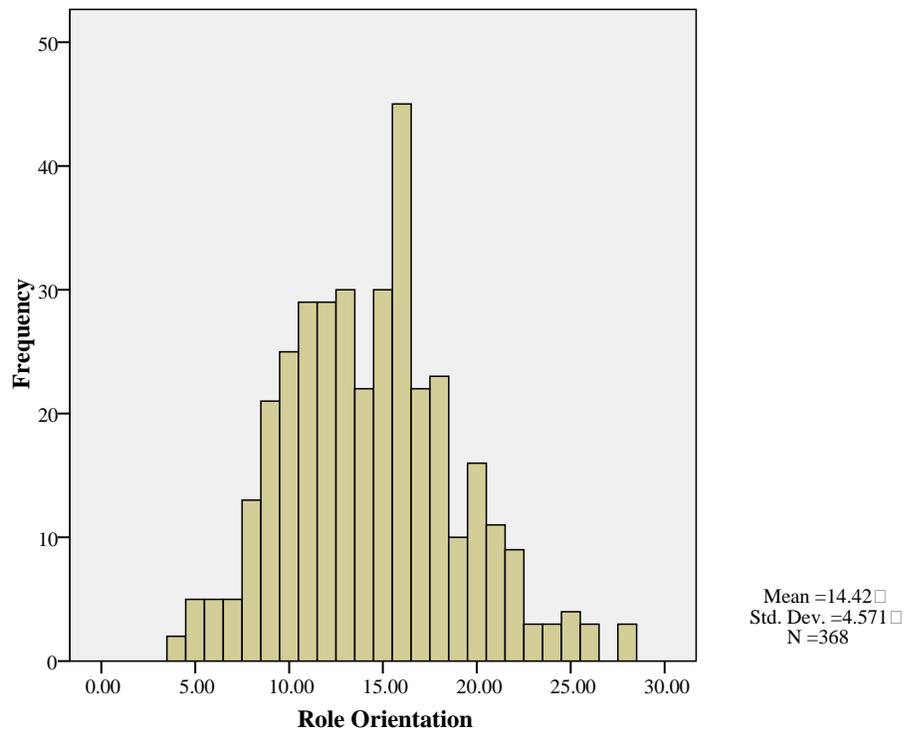
($\beta = -.09$, $SE = 0.03$, $Wald = 8.0$, $df = 1$, $p = .005$); role orientation had no impact on judicial ruling ($\beta = 0.05$, $SE = 0.04$, $Wald = 1.25$, $df = 1$, $p = .26$). Despite having a strong impact on judicial ruling, legal philosophy had no impact on dangerousness determinations ($\chi^2[6, N = 368] = 7.24$, $p = .3$), even when it was the only predictor in the model.

Hypothesis nine. The hypothesis that, as role orientation scores increased, the relation between dangerous determination and judicial ruling would strengthen was rejected. With regard to the four-item role orientation scale, participants in this study obtained scores between 4 and 28 (total range possible = 4 to 28), with the average score being 14.42 (Mode = 14, *Mdn* = 14, *SD* = 4.54; see Figure 8 for distribution of Role Orientation scores). Higher scores on this scale indicated that a judge believes that following the letter of the law is important, whereas lower scores indicated that a judge is more likely to believe discretion is appropriate to use when making judicial decisions. The internal consistency of this scale ($\alpha = .45$) was poor, which lowered the likelihood a significant effect could be detected.

The use of an ordinal regression analysis to test this hypothesis was not possible because the assumption of parallel lines was violated, $\chi^2(9, N = 355) = 37.15$, $p < .001$. As recommended by Garson (2011), in situations such as these, a multinomial logistic regression analysis should be used because it does not have restrictive assumptions. In order to decrease multicollinearity, role orientation was centered.

Figure 8

Distribution of Role Orientation Scores



When interaction effects are analyzed with multinomial logistic regressions, two models are created: one with the interaction term and the other without the interaction term. The model with the interaction term was significant ($-2LL = 559.39$, $\chi^2[12, N = 355] = 216.35$, $p < .001$, Pseudo $R^2 = .48$), as was the model without the interaction term ($-2LL = 568.41$, $\chi^2[8, N = 355] = 207.33$, $p < .001$, Pseudo $R^2 = .46$). These findings suggested the potential that the interaction effect contributed to the model. However, review of the chi square in the likelihood ratio tests indicated that the interaction contributed very little to the model, $\chi^2(4, N = 355) = 9.02$, $p = .06$. Role orientation did not predict judicial ruling ($\chi^2[4, N = 343] = 7.86$, $p = .10$), nor did it predict dangerousness determinations ($\chi^2[6, N = 347] = 8.35$, $p = .21$) when other variables (i.e., risk level, FFN, and legal philosophy) were controlled in the model; this was true even when role orientation was the only predictor in the model.

Risk level significantly influenced this potential moderation effect, $\chi^2(28, N = 355) = 43.07, p = .03$. Multinomial regression analyses indicated that there was a significant interaction effect for the low risk sample ($\chi^2[4, N = 181] = 10.7, p = .03$), but not the high risk sample ($\chi^2[4, N = 174] = 0.57, p = .97$). For the low risk sample, increases in role orientation scores were associated with less restrictive rulings ($\chi^2[1, N = 182] = 7.66, p = .006$) and less agreement that respondent was dangerousness ($\chi^2[1, N = 185] = 6.88, p = .009$). In the low risk sample, the relation between dangerousness determination and judicial ruling was stronger (Pseudo $R^2 = .32$) for judges who believed discretion is appropriate ($\chi^2[4, N = 102] = 36.58, p < .001$) than was the strength of the relation between dangerousness determination and judicial ruling (Pseudo $R^2 = .26$) for judges who believed discretion is inappropriate ($\chi^2[4, N = 81] = 22.64, p < .001$). The above mentioned finding indicated a pattern opposite of that hypothesized; specifically, judges who believe discretion is inappropriate exhibited more discordance between dangerousness and ruling, than did judges who believe discretion is appropriate. Thus, as role orientation scores increased the relation between dangerous determination and judicial ruling weakened, instead of strengthened.

CHAPTER 4

DISCUSSION

This main purpose of this study was to investigate judges' opinions regarding the probative value of different types of risk communication messages. Such information is valuable in that it can provide clinicians with tentative guidance as to which risk messages judges find most helpful when making commitment decisions. In addition, this study included an investigation into whether these risk messages impacted judicial decisions. Knowledge about whether risk messages influence judicial decisions may help clinicians express their risk assessment results in a manner that is more clear or precise. As stated by Schopp (1996), an ideal form of risk communication would provide "clear, precise, and complete information" to decision makers (p. 939).

In addition to the aforementioned practical significance of this study, this study sought to contribute theoretical information to the risk communication literature. Decades of risk assessment research have resulted in improvements in how clinicians assess risk, but research into risk communication has been relatively neglected. Scholars (e.g., Heilbrun et al., 1999; Kwartner et al., 2006; Monahan et al., 2002) have pointed out that, without risk communication, there is no link between risk assessments and decisions regarding risk. Therefore, better-informed risk decisions can only be achieved when violence risk is assessed accurately and communicated in a manner in which it can be understood by the decision maker. Considering that legal decisions regarding risk can result in the loss of individual freedoms or possibly harm to self or others, researchers have emphasized the need to improve risk communication (Heilbrun, Dvoskin, Hart & McNiel, 1999; Heilbrun et al.,

2000, 2004; Kwartner et al., 2006; Monahan et al., 2002). Despite this call for action, there has been relatively little empirical investigations into risk communication.

The Kwartner and colleagues (2006) study is the only study that investigated judges' perceptions of risk prediction messages. In many ways, the current study is an expansion of the investigation by Kwartner and colleagues (2006). The current study employed a similar methodology to that of Kwartner and colleagues (2006), but with some key differences. First, the current study consisted of a larger, national sample of judges and information about five types of risk messages. In contrast, Kwartner and colleagues (2006) focused on only risk prediction messages. In addition to examining probative value, the current study investigated the impact that these risk messages have on dangerousness determinations and commitment decisions. Because research has not been conducted on the impact risk messages have on judicial decisions, the hypotheses related to judicial rulings in the current study were based on research with samples of clinicians who were asked to make hospital discharge decisions (e.g., Monahan et al., 2002; Slovic & Monahan, 1995; Slovic et al., 2000).

Although the primary focus of the current study was on risk communication, this study also sought to investigate the impact that judicial attributes may have on the judicial decision-making process. Despite conjecture that judges only consider legal factors when making decisions (e.g., Corwin, 1924; Mendelson, 1963; Schauer, 1988), research has indicated that the judicial decision-making process may consist of a complex schema that includes the influence of extralegal factors, such as role orientation and sentencing goals (e.g., Gibson, 1978; Homel & Lawrence, 1992). The current study investigated the impact that three judicial attributes, or extralegal factors, had on the decision-making process: role orientation, legal philosophy, and FFN. The investigation of these factors is important because it can provide insight into whether commitment decisions or dangerous determinations are impacted, or moderated, by judicial attributes.

Probative Value of Risk Messages

As previously mentioned, the current study was an expansion of the work of Kwartner and colleagues (2006). Similar to that study, results of the current study indicated that judges rated categorical messages as the risk prediction format with the highest probative value. Kwartner and colleagues (2006) postulated that judges may rate categorical messages as most probative because they are “the easiest to understand and interpret” (p. 192). This rationale was tentatively supported by the finding, in the current study, that numerical estimates tended to result in more inaccurate commitment decisions than did categorical messages.

The current study also found that judges rated the probative value of any risk message higher in the high risk scenarios, as compared to the low risk scenarios. Researchers have found similar results with a sample of judges (Kwartner et al., 2006) and with samples of clinicians (Heilbrun et al., 2000, 2004). In high risk scenarios, any information, regardless of how it is communicated, aids judges in the decision-making process. Overall, judges assigned the highest probative value to the high risk vignette that included a categorical prediction message. Due to the more complex nature of high risk scenarios, the need for an easily understood risk message is of even greater importance.

Although there was a significant difference between probative value ratings for risk prediction formats, judges did not perceive any risk model as significantly more probative in value. However, judges exhibited a slight tendency to rate risk prediction and management messages as more probative in value than description messages. The initial hypothesis that judges would prefer the risk management model was based on research that has indicated clinicians prefer risk management messages to risk prediction or description messages (Heilbrun et al., 2000; 2004).

It is possible the results of the current study were discordant with the Heilbrun and colleagues’ (2000, 2004) studies because clinicians and judges may have fundamentally different

opinions about what type of information is probative. For example, Heilbrun and colleagues (2000; 2004) postulated that clinicians preferred risk management messages because prediction statements were more likely to be viewed as prejudicial and were more likely to result in a “premature closure on the deliberation of the decision of the decision maker” (Heilbrun et al., 2000, p. 145; Heilbrun et al., 2004, p. 195). In addition, Heilbrun and colleagues (2000) found that clinicians who primarily conducted research were less likely to rate risk management messages as highly probative, than were clinicians who primarily involved in clinical or forensic practice. Thus, clinicians who provide treatment may consider the extent to which treatment may help the respondent as more important than the respondent’s predicted risk level. It is possible that the opinion of judges is more aligned with the opinion of research clinicians than the opinion of clinicians who are involved in practice.

Results of the current study indicated judges found risk prediction statements just as probative in value as intervention suggestions. This finding suggested that judges do not perceive prediction messages as prejudicial. Because a goal of risk communication research is to provide complete information to decision makers, clinicians may want to provide both prediction and management information to judges. Despite the finding that the clinicians prefer risk management messages, Heilbrun and colleagues (2000; 2004) also suggested that clinicians provide “both predictive and management-oriented risk information” (Heilbrun et al., 2000, p. 146). Thus, the results of the current study provided further support for this type of approach, in which clinicians provide both risk prediction and management information.

Impact of Risk Messages on Judicial Ruling

Contrary to hypothesis three, risk prediction format did not affect restrictiveness of judicial ruling. Specifically, it was hypothesized that frequency terms would result in the most restrictive rulings. This hypothesis was based on research that asked clinicians to decide if they would discharge

a respondent from a hospital based on information provided through frequency or probability risk messages (Monahan et al., 2002; Slovic et al., 2000). In these studies, forensic clinicians were more likely to decline release if the clinician was given the information in a frequency message. The authors postulated that frequency messages may be “fear-inducing risk communication formats” and that those with personal experiences with false negative may be most susceptible to the affects of frequency messages. Results from the current study indicated that frequency statements did not induce fear in the judges, even when looking specifically at judges who tended to fear false negative errors and those who were aware that they had made false negative errors in the past.

There are two possible reasons that the findings related to risk prediction format were discrepant with prior results in the literature (Monahan et al., 2002; Slovic et al., 2000). Similar to clinicians’ and judges’ differences regarding the information they find probative, their decision-making process may differ to such an extent that results based on samples of forensic clinicians cannot be extrapolated to judicial samples. For example, Monahan and colleagues (2002) found that for clinicians who worked in nonforensic settings, frequency statements did not impact the restrictiveness of their decisions. These authors postulated that forensic clinicians may be more influenced by frequency statements, because they have more experience with false negative predictions and the impact these errors have on the victim of the respondent’s violence. It is possible that judges are more similar to nonforensic clinicians than they are to forensic clinicians. However, this study did not assess similarities between judges and nonforensic clinicians and therefore it is unknown what characteristics judges and nonforensic clinicians may share that makes them less likely to be influenced by frequency statements.

The discrepant results between the current study and the aforementioned studies (Monahan et al., 2002; Slovic et al., 2000) may also be due to the differences in the operational definition of

commitment decisions. In both of the aforementioned studies, participants were asked to make decisions regarding an individual that was already committed to a psychiatric facility. It is possible the decision-making process in recommitment hearings differs substantially from the decision-making process in initial commitment hearings. Although Parry, Turkheimer, and Hundley (1992) found that there were no significant differences regarding the presentation of evidence supporting dangerousness between these two hearings, it is possible that other elements differ substantially enough to make the extrapolation inappropriate.

Although the results of this study did not confirm the hypothesis that risk prediction format impacted judicial ruling, there was tentative evidence that risk messages have an impact on judicial ruling. Specifically, this study found the judges who received risk prediction messages were more likely to make restrictive rulings than were judges who received risk description or management messages. The current study was the first to analyze the impact risk models have on commitment decisions. Therefore, further research is needed to replicate these findings. If future studies confirm this result, it may be possible the risk prediction messages, in general, cause judges to visualize false negatives and result in a tendency to err on the side of caution.

Although risk prediction format did not impact judicial rulings, the results of this study indicated that risk prediction format may influence the judicial decision-making process in a manner not previously demonstrated in the literature. Risk prediction format impacted dangerousness differently depending on the risk level. Judges who received frequency messages were most likely to decide a low risk respondent was dangerous, but least likely to decide a high risk respondent was dangerous. Because risk prediction format had no significant impact on judicial ruling in either the low risk or high risk conditions, it cannot be concluded there was a significant mediation effect. However, results indicated there was an indirect effect when examining the low risk sample: in the

low risk sample, risk prediction impacted dangerousness determinations, which in turn impacted judicial rulings.

Although the results regarding the impact of frequency messages on judicial ruling were discordant with the findings of Monahan et al. (2002) and Slovic et al. (2000), the results of the current study are in line with conjectures provided by Heilbrun and colleagues (1999) and Monahan and Steadman (1996). In these two articles, the authors presented the argument that numerical estimates are more easily misinterpreted than are categorical messages. The results of the current study supported this argument when investigating the accuracy of commitment decisions, instead of ruling restrictiveness. Compared to numerical estimates, categorical messages resulted in more accurate commitment decisions, with regard to sensitivity, specificity, NPP, PPP, and overall correct classification.

Inaccurate commitment decisions can have serious consequences either on the liberty interests of the respondent or on the safety of the community. For example, numerical estimates resulted in lower sensitivity, which means that slightly more judges who received the numerical messages erroneously decided to permit a dangerous respondent to remain in the community. This false negative error could potentially result in community members being harmed by the respondent. Although minor violent behaviors (e.g., destruction of property; slapping; pushing) are more common, serious violent behaviors could have a devastating impact on a victim and could potentially impact the credibility of the judge. In order to avoid these negative consequences, judges strive to make accurate commitment decisions.

The Impact of Judicial Attributes on Legal Decisions

Some scholars have opined that judges only consider legal factors and do not allow extralegal factors to impact their decisions (e.g., Brisbin, 1996; Schauer, 1988). The results of this study

indicated that extralegal factors predicted ruling in a manner that was above and beyond that which was predicted by the objective legal factor of risk level. A judge's opinion about the purpose of his or her ruling (i.e., legal philosophy) and about the consequences of his or her error (i.e., FFN) had a significant impact on his or her judicial ruling. A judge's opinion about his or her role in the judiciary (i.e., role orientation) did not have such an impact.

Legal philosophy.

With regard to legal philosophy, results indicated that judges with *parens patria* stances were more likely to make a restrictive ruling than their police power counterparts. This finding supported Hiday's (1983) postulation that judges who endorse a *parens patria* stance are more likely to commit a respondent. Although legal philosophy had no direct impact on dangerousness determinations, legal philosophy impacted the strength of the relation between dangerousness determinations and judicial ruling. This finding suggests that legal philosophy affects the manner in which judges make their commitment decisions. According to the results of this study, judges with police power stance exhibited concordance between their decisions regarding dangerousness and judicial ruling, while *parens patria* judges were more likely to exhibit discordance between these two decisions. Thus, judges with police power philosophies decided that the danger to others criterion was present and of sufficient strength to warrant commitment. In contrast, *parens patria* judges were more likely to decide that the presence of danger to others was not necessary to warrant commitment.

The possibility that some judges disregard the criterion of danger to others is in line with research that has indicated that judges tend to consider *parens patria* notions more strongly than police power notions. In Bursztajn and colleagues' (1986; 1997) studies, judges indicated that danger to self predicted ruling more than did danger to others. The current study did not include a reference to danger to self, because inclusion would have confounded the results of the study by restricting a

judge's option or by making the need to consider the danger to others criterion obsolete. Although judges were told to base their commitment decisions exclusively on their dangerousness determinations, it is possible that some judges considered *parens patriae* notions (e.g., the family desired commitment; assumptions about the respondent's danger to self) when making commitment decisions.

Judges who believe it is their role to care for mentally ill individuals may rule that commitment is necessary, even if the individual is not dangerous. In *O'Connor v. Donaldson* (1975), the justices opined that the mere presence of mental illness "cannot justify a State's locking a person up against his will" (p. 575). For this reason, in order to justify commitment, all civil commitment statutes require that a respondent be mentally ill and be a danger to self or others. Most experienced judges are familiar with the applicable criteria and recognize that each criterion must be objectively considered in each case (Bursztajn et al., 1997). In order to increase objectivity, a judge should be aware of the factors they intuitively rely on, and the potential pitfall associated with the identified factors. In this study, there was no presented evidence that the respondent was a danger to self; therefore, judges who justified commitment on the basis of danger to self either believe medication noncompliance or substance abuse always equate to self-harm or they assume that evidence for self-harm is usually present, even when not mentioned. By recognizing that legal philosophy impacts their decisions, judges may be better able to analyze whether they are objectively applying the law and may be better able to avoid the pitfalls associated with the factors they intuitively consider.

Fear of false negative (FFN).

FFN had a different impact on the judicial decision-making process than did legal philosophy. Compared to their low FFN counterparts, judges with high FFN were significantly more likely to make a more restrictive ruling and more likely to opine that the respondent was dangerous. Results

indicated that FFN did not moderate the relation between risk prediction format and judicial ruling. In this situation, a moderation effect would have been difficult to find, because results indicated there was no direct association between risk prediction format and judicial ruling. However, dangerousness and judicial ruling were strongly related, and FFN did not moderate their relation. Although the proposed moderation hypothesis was rejected, the findings of the current study are in line with prior researchers/scholars (e.g., Durham, 1999; Heilbrun, 1997; Wexler, 1983; Winick, 2005) that have suggested that judges with high FFN tend to err on the side of caution and make restrictive decisions.

In conclusion, judges with high FFN were more likely to agree that dangerousness was present and were more likely to find that the level of dangerousness was sufficient to warrant commitment. For judges with high FFN, this tendency to err on the side of caution may result in false positive errors. The results of this study indicated that judges with high FFN exhibited lower specificity (.27) than did judges with low FFN (.86). Thus, a large proportion of high FFN judges erroneously decided to deprive a nondangerous respondent of his civil liberties. As the result of this type of error, a respondent would be subject to the stigma associated with placement in a psychiatric facility and to the potential financial loss associated with treatment costs or inability to maintain employment (*Goetz v. Crosson*, 1992). Similar to the results regarding legal philosophy, knowledge about the impact of FFN may allow judges to analyze the objectivity of their decisions and personally develop ways to decrease any potential subjectivity.

Role orientation.

Role orientation was the only judicial attribute that did not significantly impact judicial ruling or dangerousness determinations. There was no moderating impact of role orientation, on the relation between dangerousness and ruling, but there was a slight indication of a pattern opposite of that hypothesized. That is, judges who believed discretion is inappropriate were more likely to decide that

a low risk respondent needed inpatient commitment. The reasons behind these results are unclear, but it is possible this scale was measuring a different construct than role orientation. The items in this scale were based on research that investigated the concept of role orientation by dividing it into different typologies (e.g., law interpreters versus activists). Extrapolating items from those research measures to create a continuous role orientation scale may have been inappropriate.

Limitations

Although the results of the current study provided information about risk communication and judicial attributes, some of the results did not match what was hypothesized. The discrepancy between obtained results and proposed hypotheses may be due to the limitations of this study. One potentially detrimental limitation of this study is the low reliability of the judicial attributes scales. Although these scales were based on statements in the research literature and pilot tested, the low reliability of the scales require the reader to practice caution when interpreting any of the findings related to judicial attributes. It is possible that the scales are not truly representing the constructs of role orientation, legal philosophy, and FFN.

Another potential limitation is the lack of ecological validity of this study. Real civil commitment cases are more complex than the hypothetical vignettes presented to the judges in this study. Judges usually make multiple decisions about the civil commitment criteria, unlike this study where judges were asked to focus exclusively on the danger to others criterion. In addition, clinicians report an array of information about the respondent that is far more extensive than the information provided in the one paragraph describing the hypothetical respondent.

Finally, it is acknowledged that the response rate was low; although it was in line with prior research using judicial samples (e.g., Kwartner et al., 2006; Manuto & O'Rourke, 1991). The sample size of 403 allowed for a precision of 4.75 when making inferences to the true population. Although

the inference precision was adequate, it is possible that judges who responded differed from those who chose not to respond. Strategies were utilized in this study to decrease the likelihood of this difference. A comparison between judges who participated after first contact versus those who participated after second contact indicated there were no significant differences between the two groups. Similar comparisons have been used in previous research to investigate the difference between responders and nonresponders. Although the overall sample size was adequate, the regional representation was more limited. The regional sample sizes did not allow for precise inferences and as such, analyses into regional differences were not presented.

Recommendation for Future Research

The current study focused on judge's probative opinions, dangerousness determinations, and commitment decisions. It is recommended that future studies obtain knowledge regarding judges' confidence in, and perceived need of, clinical risk communications in commitment decisions as this information would be helpful in understanding opinions about risk messages. Similarly, it may be beneficial for future studies to gather empirical information regarding the interplay between the civil commitment criteria of danger to others, danger to self decisions, and mental illness. Although the regional sample size of this study was too small to make inferences to the population, it would be informative if future studies investigated how civil commitment statutes and judges differ by region. Investigations into regional difference can help clinicians understand which risk communication findings are applicable to them. Thus, it is recommended that the sample size for future studies be large enough to make precise inferences to the regional population.

Conclusion

As previously mentioned, this study sought to answer eight research questions regarding decision making in civil commitment proceedings. The most important question pertained to

identifying which type of risk message judges find most probative in value. The study found that no risk model was deemed most probative, but there was a slight, nonsignificant, tendency to rate as more probative, messages that included explicit conclusions regarding predictions or treatment recommendations. When utilizing risk prediction messages, clinicians may want to express their predictions via categorical terms, because judges found these messages to be the most probative in value.

Another important research question addressed the impact of risk messages on judicial rulings. Results of the current study indicated that risk prediction messages, in general, led to stricter rulings than did risk management or description messages. Within risk prediction formats, no format (i.e., categorical, probabilistic, or frequency) resulted in higher restrictiveness in judicial ruling. Through an investigation into the accuracy of commitment decisions, this study found that categorical messages, as compared to numerical estimates, resulted in more accurate decisions, with regard to sensitivity, specificity, NPP, PPP, and overall correct classification. This finding suggested that categorical statements are easier to interpret than are numerical statements.

This study was the first study to investigate the impact that judicial attributes have on the decision-making process in civil commitment proceedings. These results, while tentative, indicated that the extralegal factors of legal philosophy and FFN directly impacted judicial rulings. In particular, certain types of judges (i.e., those with a *parens patriae* stance or high FFN) were more likely to make restrictive rulings than were their counterparts. A judge's opinion about the purpose of civil commitments had a significant impact on the information they considered when making a commitment decision. Specifically, judges who believed that civil commitments were in the respondent's best interests and necessary to benefit his or her health tended to make more restrictive rulings despite not finding evidence the respondent was dangerous.

In conclusion, this study sought to provide information that could assist clinicians in communicating their risk assessments in a clear, concise, and complete manner. The secondary purpose of this study was to contribute theoretical information to the risk communication and judicial decision-making literature. The fundamental goal of this study was to provide information that can help judges make better-informed legal decisions.

REFERENCES

- Addington v. Texas, 441 U.S. 418 (1979).
- Allbright, A., Levy, F., & Wagle, N.C. (2002). Outpatient civil commitment laws: An overview. *Mental and Physical Disability Law Reporter*, 26, 179-185.
- The American Bench. Judges of the Nation.* (2010). Sacramento: Forster-Long
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (text rev.). Washington, DC: Author.
- Anderson, R., Pichert, J.W., & Shirey, L.L. (1983). Effects of the reader's schema at different points in time. *Journal of Educational Psychology*, 75, 271-279.
- Appelbaum, P. (1992). Civil commitment from a systems perspective. *Law and Human Behavior*, 16, 61-74.
- Baker, E. (1995). Public response to hurricane probability forecasts. *Professional Geographer*, 47, 137-147.
- Baron, R., & Kenny, D. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Bartels, S.J., Drake, R.E., Wallach, M.A., & Freeman, D.H. (1991). Characteristic hostility in schizophrenic outpatients. *Schizophrenia Bulletin*, 17, 163-171.
- Baum, L. (1997). *The puzzle of judicial behavior*. Ann Arbor, MI: University of Michigan Press.
- Blanck, P.D., Rosenthal, R., Hart, A.J., & Bernieritt, F. (1989). The measure of the judge: An empirically-based framework for exploring trial judges' behavior. *Iowa Law Review*, 75, 653-684.
- BNA's Directory of State and Federal Courts, Judges, and Clerks.* (2010). Washington, D.C.: Bureau of National Affairs.
- Borum, R., & Verhaagen, D. (2006). *Assessing and managing violence risk in juveniles*. New York: Guildford Press.
- Brisbin, R. (1996). Slaying the dragon: Segal, Speath, and the function of law in Supreme Court decision making. *American Journal of Political Science*, 40, 1004-1017.
- Brooks, R. (2007). Psychiatrists' opinions about involuntary civil commitment: Results of a national survey. *Journal of American Academy of Psychiatry and Law*, 35, 219-228.

- Brown v. Board of Education of Topeka, 347 U.S. 483 (1954).
- Bursztajn, H., Gutheil, T.G., Mills, M., Hamm, R.M., & Brodsky, A. (1986). Process analysis of judges' commitment decisions: A preliminary empirical study, *American Journal of Psychiatry*, *143*, 170-174.
- Bursztajn, H., Hamm, R.M., & Gutheil, T.G. (1997). Beyond the black letter of the law: An empirical study of an individual judge's decision process for civil commitment hearings. *Journal of the American Academy of Psychiatry and the Law*, *25*, 79-94.
- Carroll, J.S., Perkwitz, W.T., Lurigio, A.J., & Weaver, F.M. (1987). Sentencing goals, causal attributions, ideology, and personality. . *Journal of Personality and Social Psychology*, *52*, 107-118.
- Clancy, K., Bartolomeo, J., Richardson, D., & Wellford, C. (1981). Sentence decision making: The logic of sentence decisions and the extent and sources of sentence. *The Journal of Criminal Law and Criminology*, *72*, 524-554.
- Conroy, M. A., & Murrie, D. C. (2007). *Forensic assessment of violence risk: A guide for risk assessment and risk management*. Hoboken, NJ: John Wiley & Sons, Inc.
- Cook, C., Heath, F., & Thompson, R.L. (2000). A meta-analysis of response rates in web- or internet-based surveys. *Educational and Psychological Measurement*, *60*, 821-836.
- Corwin, E. (1924). The constitutional decisions of the Supreme Court of the United States in the October term. *The American Political Science Review*, *18*, 49-78.
- Davis, T. L., Severy, L. J., Kraus, S. J., & Whitaker, J. M. (1993). Predictors of sentencing decisions: The beliefs, personality variables, and demographic factors of juvenile justice personnel. *Journal of Applied Social Psychology*, *23*, 451-477.
- DeCoster, J. (1998). *Overview of Factor Analysis*. Retrieved October 2, 2011 from <http://www.stat-help.com/notes.html>
- DeCoster, J. (2006). *Testing Group Differences using T-tests, ANOVA, and Nonparametric Measures*. Retrieved March 23, 2007 from <http://www.stat-help.com/notes.html>
- DeCoster, J. (2007). *Applied Linear Regression Notes set 1*. Retrieved November 11, 2007 from <http://www.stat-help.com/notes.html>
- DeCoster, J., & Claypool, H. (2004). *Data Analysis in SPSS*. Retrieved July 24, 2010 from <http://www.stat-help.com/notes.html>
- Dillman, D.A., Smyth, J.D., & Christian, L.M. (2009). *Internet, mail and mixed-mode surveys: The tailored design method*. Hoboken, NJ: John Wiley & Sons, Inc.

- The Directory of Minority Judges of the United States.* (2008). Chicago, Ill: American Bar Association.
- Dixon v. Attorney General of the Commonwealth of Pennsylvania, 325 F.Supp. 966, 972, 974 (M.D.Pa.1971).
- Douglas, K.S., & Skeem, J.L. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy and Law*, 3, 347-383.
- Durham, M.L. (1996). Civil commitment of the mentally ill: Research, policy and practice. In B. Sales, & S. Shah (Eds.), *Mental health and law: Research, policy and services* (pp. 17-40). Durham, N.C.: Carolina Academic Press.
- Fariña, F., Arce, R., & Novo, M. (2003). Anchoring in judicial decision-making. *Psychology in Spain*, 7, 56-65.
- Fiske, S.T., & Taylor, S.E. (1991). *Social Cognition*. New York: McGraw-Hill.
- Frank, J. (1949). *Courts on Trial: Myth and Reality in American Justice*. Princeton, NJ: Princeton University Press.
- Garson, J.D. (2009). *Sampling*. Retrieved August 20, 2010 from <http://faculty.chass.ncsu.edu/garson/PA765/sampling.htm>
- Garson, J.D. (2010). *Testing of Assumptions*. Retrieved July 13, 2010 from <http://faculty.chass.ncsu.edu/garson/PA765/assumpt.htm#normal>
- Gibson, J.L. (1977). Discriminant functions, role orientations and judicial behavior: Theoretical and methodological linkages. *The Journal of Politics*, 39, 984-1007.
- Gibson, J.L. (1978). Judges' role orientations, attitudes, and decisions: An interactive model. *American Political Science Review*, 72, 911-924.
- Gibson, J.L. (1981). The role concept in judicial research. *Law and Policy Quarterly*, 3, 291-311.
- Gibson, J.L. (1983). From simplicity to complexity: The development of theory in the study of judicial behavior. *Political Behavior*, 5, 7-49.
- Gillman, H. (2001) What's law got to do with it? Judicial behavioralists test the 'legal model' of judicial decision making. *Law and Social Inquiry*, 465-504.
- Goetz v. Crosson, 967 F.2d 29 (1992).
- Goritz, A., & Stieger, S. (2009). The impact of the field time on response, retention, and response completeness in list-based Web surveys. *International Journal of Human-Computer Studies*, 67, 342-348.

- Grisso, T., & Tomkins, A. (1996). Communicating Violence Risk Assessments. *American Psychologist*, 51, 928-930.
- Groves, R.M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation. *Public Opinion Quarterly*, 64, 299-308.
- Guthrie, C., Rachlinski, J.J., & Wistrich, A. (2001). Inside the judicial mind. *Cornell Law Review*, 86, 777-830.
- Heilbrun, K. (1997). Prediction versus management models relevant to risk assessment: The importance of legal decision making context. *Law and Human Behavior*, 21, 347-359.
- Heilbrun, K., Dvoskin, J., Hart, S., & McNeil, D. (1999). Violence risk communication: Implications for research, policy, and practice. *Health, Risk & Society*, 1, 91-106.
- Heilbrun, K., O'Neill, M. L., Strohman, L. K., Bowman, Q., & Philipson, J. (2000). Expert approaches to communicating violence risk. *Law and Human Behavior*, 24, 137-148
- Heilbrun, K., O'Neill, M.L., Stevens, M.A., Strohman, L.K., Bowman, Q., & Lo, Y. (2004). Assessing normative approaches to communicating violence risk: National survey of psychologists. *Behavioral Sciences and the Law*, 22, 187-96.
- Heilbrun, K., Philipson, J., Berman, L., & Warren, J. (1999). Risk communication: Clinician's reported approaches to communicating violence risk. *Law and Human Behavior*, 27, 397-406.
- Heise, M. (2002). The past, present, and future of empirical legal scholarship: Judicial decision making and the New Empiricism, *University of Illinois Law Review*. 819-850.
- Herrenkohl, T.I., Hill, K.G., Chung, I.J., Guo, J., Abbott, R., & Hawkins, J.D. (2003). Protective factors against serious violent behavior in adolescence: A prospective study of aggressive children. *Social Work Research*, 27, 179-191.
- Hiday, V.A. (1983). Judicial decisions in civil commitment: Facts, attitudes, and psychiatric recommendations. *Law & Society Review*, 17, 517-529.
- Hilton, N.Z., Harris, G.T., Rawson, K., & Beach, C.A. (2005). Communicating violence risk information to forensic decision makers. *Criminal Justice and Behavior*, 32, 97-116.
- Homel, R. J., & Lawrence, J. A. (1992). Sentencer orientation and case details: An interactive analysis. *Law and Human Behavior*, 16, 509-537.
- Howard, J.W. (1977). Role perceptions and behavior in three US Courts of Appeals. *The Journal of Politics*, 39, 916-938.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.

- Kulik, C.T., Perry, E.L., & Pepper, M.B. (2003). Here comes the judge: The influence of judge personal characteristics on federal sexual harassment case outcomes. *Law and Human Behavior, 27*, 69-86.
- Kwartner, P., Lyons, P., & Boccaccini, M. (2006). Judges' risk communication preferences in risk for future violence cases. *International Journal of Forensic Mental Health, 5*, 185-194.
- Lessard v. Schmidt, 349 F. Supp. 1078 (1972).
- Levi, E. (1948). An introduction to legal reasoning. *The University of Chicago Law Review, 15*, 501-574
- Litwark, T. (2002). Some questions for the field of violence risk assessment and forensic mental health: Or, "Back To Basics" revisited. *International Journal of Forensic Mental Health, 1*, 171-178.
- Llewellyn, K. (1930). A realistic jurisprudence – The next step. *Columbia Law Review, 30*, 431-465.
- Lodewijks, H.P.B., Ruiter, C., & Doreleijers, T.A.H. (2010). The impact of protective factors in desistance from violent reoffending: A study in three samples of adolescent offenders. *Journal of Interpersonal Violence, 25*, 568-587.
- McNiel, D., & Binder, R. (1998, March). *Comparison of categorical and probabilistic approaches to communication about psychiatric patients' risk of violence in clinical practice*. Paper presented at the Biennial Conference of the American Psychological-Law Society/Division 41 of the American Psychological Association, Redondo Beach, CA.
- Melton, G. B., Petrila, J., Poythress, N., & Slobogin, C. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers*. New York: Guilford Press.
- Mendelson, W. (1963). The neo-behavioral approach to the judicial process: A critique. *The American Political Science Review, 57*, 593-603.
- Microsoft Corporation (2010). *Use mail merge to send personalized e-mail messages to your e-mail address list*. Retrieved September 2, 2010 from <http://office.microsoft.com/en-us/word-help/use-mail-merge-to-send-personalized-e-mail-messages-to-your-e-mail-address-list-HA010109162.aspx>
- Monahan, J. (1981). *The clinical prediction of violent behavior*. Rockville, MD: National Institute of Mental Health.
- Monahan, J., Heilbrun, K., Silver, E., Nabors, E., Bone, J., & Slovic, P. (2002). Communicating violence risk: Frequency formats, vivid outcomes, and forensic settings. *International Journal of Forensic Mental Health, 1*, 121-126.

- Monahan, J., & Silver, E. (2003). Judicial decision thresholds for violence risk management. *International Journal of Forensic Mental Health, 2*, 1-6.
- Monahan, J., & Steadman, H. (1996). Violent storms and violent people: How meteorology can inform risk communication in mental health law. *American Psychology, 51*, 931-938.
- Monahan, J., Steadman, H., Silver, E., Appelbaum, P.S., Robbins, P.C., Mulvey, E.P., et al. (2001). *Rethinking risk assessment: the MacArthur study of mental disorder and violence*. New York: Oxford University Press.
- Monahan, J., & Wexler, D.B. (1978). A definite maybe: Proof and probability in civil commitment. *Law and Human Behavior, 2*, 37-42.
- Manuto, R., & O'Rourke, S.P. (1991). Federal judges' perceptions of social research in judicial decision-making. *Communication Reports, 4*, 103-106.
- Morrison, G.M., Robertson, L., Laurie, B., & Kelly, J. (2002). Protective factors related to antisocial behavior trajectories. *Journal of Clinical Psychology, 58*, 277 – 290.
- Mrad, D., & Nabors, E. (2006). The role of the psychologist in civil commitment. In A. Goldstein (Ed.), *Forensic psychology: Emerging topics and expanding roles* (pp. 232-259). Hoboken, NJ: John Wiley & Sons, Inc.
- Mulvey, E.P., & Lidz, C.W. (1985). Back to basics: A critical analysis of dangerousness research in a new legal environment. *Law and Human Behavior, 9*, 209-219.
- National Center for State Court (2009) *Probate Courts*. Retrieved from the World Wide Web on June 16, 2009 <http://www.ncsconline.org/WC/CourTopics/statelinks.asp?id=69&topic=ProbCt>
- O'Conner & Donaldson, 422 U.S. 563 (1975).
- Pallant, J. (2007). *SPSS survival manual. A step by step guide to data analysis using SPSS for Windows*. Buckingham: Open University Press.
- Parry, J. (1994). Involuntary civil commitment in the 90s: A constitutional perspective. *Mental and Physical Disability Law Reporter, 18*, 320-336.
- Parry, C.D., Turkheimer, E., & Hundley, P.L. (1992). A comparison of commitment and recommitment hearings: Legal and policy implications. *International Journal of Law and Psychiatry, 15*, 25-41.
- Posner, R.A. (2006). *The Role of the Judge in the Twenty-First Century*. Boston University Law Review, 86, 1049-1068.
- Rogers, R. (2000). The uncritical acceptance of risk assessment in forensic practice. *Law and Human Behavior, 24*, 595-605.

- Rowland, C.K., & Carp, R. (1996). *Politics and Judgment in Federal District Courts*. Lawrence, KS: University Press of Kansas.
- Salant, P., & Dillman, D.A. (1994). *How to Conduct Your Own Survey*. New York: John Wiley and Sons.
- Schauer, F. (1988). Formalism. *The Yale Law Journal*, 97, 509-548.
- Scheb, J.M., Bowen, T., & Anderson, G. (1991). Ideology, Role Orientations, and Behavior in the State Courts. *American Politics Research*, 19, 324-335.
- Scheb, J.M., & Unga, T.D. (1986). Competing orientations to the judicial role: The case of Tennessee judges. *Tennessee Law Review*, 54, 391-412.
- Scheb, J.M., Unga, T.D., & Hayes, A.L. (1989). Judicial role orientations, attitudes and decision making: A research note. *Political Research Quarterly*, 42, 427-435.
- Schopp, R.F. (1996). Communicating risk assessment Accuracy, efficacy and responsibility. *American Psychology*, 51, 931-938.
- Schopp, R.F. (2001). *Competence, Condemnation, and Commitment: An Integrated Theory of Mental Health Law (Law and Public Policy: Psychology and the Social Sciences)*. Washington, DC: American Psychological Association.
- Scott, C., & Resnick, P. (2006). Violence risk assessment in persons with mental illness. *Aggression and Violent Behavior*, 11, 598-611.
- Segal, J. & Spaeth, H. (1996). The influence of stare decisis on the votes of the United States Supreme Court justices. *American Journal of Political Science*, 40, 971-1003.
- Sheehan, K. B., & Hoy, M. B. (1999). Using e-mail to survey internet users in the United States: Methodology and assessment. *Journal of Computer Mediated Communication*, 6. Retrieved November 19, 2009, from <http://jcmc.indiana.edu/vol6/issue2/sheehan.html>
- Simon, D. (1998). A Psychological Model of Judicial Decision Making. *Rutgers Law Journal*, 30, 1-142
- Slovenko, R. (2006). Remorse commentary. *Journal of Psychiatry and Law*, 34, 397-432.
- Slovic, P., & Monahan, J. (1995). Probability, danger, and coercion: A study of risk perception and decision making in mental health law. *Law and Human Behavior*, 19, 49-65.
- Slovic, P., Monahan, J., & MacGregor, D.G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior*, 24, 271-296

- Smith, C.E. (1993). *Courts, politics, and the judicial process*. Chicago, IL: Nelson-Hall Publishers.
- Steadman, H. (1987). How well can we predict violence for adults? In F. Dutilleul & C. Foust (Eds.). *The prediction of criminal violence* (pp. 5-16). Springfield, IL: Charles C Thomas.
- Steadman, H., Monahan, J., Appelbaum, P., Grisso, T., Mulvey, E., Roth, L., Robbins, P., & Klassen, D. (1994). Designing a new generation of risk assessment research. In J. Monahan & H. Steadman (Eds.), *Violence and Mental Disorder: Developments in Risk Assessment* (pp. 297 – 318). Chicago: University of Chicago Press. Retrieved July 2, 2010 at <http://books.google.com/books?hl=en&lr=&id=v4oSb2Na4nQC&oi>
- Steadman, H., Silver, E., Monahan, J., Appelbaum, P. S., Robbin, P. C., Mulvey, E. P., et al. (2000). A classification tree approach to the development of actuarial violence risk assessment tools. *Law and Human Behavior*, 24, 83-100.
- Survey Monkey. (2007). *Survey Monkey User Manual: User Guide for Managing Accounts, Creating, Distributing, & Analyzing Surveys within SurveyMonkey*. Retrieved from the World Wide Web on June 13, 2009 <http://s3.amazonaws.com/SurveyMonkeyFiles/UserManual.pdf>
- Swanson, J.W., Holzer, C.E., Ganju, V.K., & Jono, R.T. (1990). Violence and psychiatric disorder in the community: Evidence from the epidemiologic catchment area surveys. *Hospital and Community Psychiatry*, 41, 761-770.
- Swanson, J.W., Swartz, M.S., Borum, R., Hiday, V.A., Wagner, R., & Burns, B.J. (2000). Involuntary out-patient commitment and reduction of violent behaviour in persons with severe mental illness. *British Journal of Psychiatry*, 176, 324 – 331.
- Swartz, M.S., Swanson, J.W., Kim, M., & Pertila, J. (2006). Use of outpatient commitment or related civil commitment treatment orders in five U.S. communities. *Psychiatric Services*, 57, 343-349.
- Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics*. Boston : Pearson/Allyn & Bacon.
- Turkheimer, E., & Parry, C.D.H. (1992). Why the Gap? Practice and Policy in Civil Commitment Hearings. *American Psychologist*, 47, 646-655.
- US Census Bureau. (2009). *Census Regions and Divisions of the United States*. Retrieved November 25, 2009, from http://www.census.gov/geo/www/us_regdiv.pdf
- Ungs, T.D., & Baas, L.R. (1972). Judicial role perceptions: A Q-technique study of Ohio judges. *Law and Society Review*, 6, 343-366.
- Vance, J.E., Bowen, N.K., Fernandez, G., & Thompson, S. (2002). Risk and protective factors as predictors of outcome in adolescents with psychiatric disorder and aggression. *Journal of American Academy of Child and Adolescent Psychiatry*, 41, 1 36- 1 43.

- Van Koppen, P.J., & Kate, J.T. (1984). Individual differences in judicial behavior: Personal characteristics and private law decision-making. *Law and Society Review*, 18, 225-248.
- Viscusi, W.K. (1999). How do judges think about risk?. *American Law and Economics Review*, 1, 26-62.
- Volavka, J., & Tardiff, K. (1999) Substance abuse and violence. In K. Tardiff (Ed.), *Medical management of the violent patient: clinical assessment and therapy* (pp. 153-171). New York: Informa HealthCare. Retrieved August 28, 2009 at <http://books.google.com/books?hl=en&lr=&id=SEKuPP20eLsC&o>
- Von Helversen, B., & Rieskamp, J. (2009). Predicting sentencing for low-level crimes: Comparing models of human judgment. *Journal of Experimental Psychology: Applied*, 15, 375-395.
- Ward, B. (2006). Sentencing without remorse, *Loyola University Chicago Law Journal*, 38, 131-167.
- Werner, E.E. (1995). Resilience in development. *Current directions in psychological science*, 4, 81-85.
- Wexler, D. (1983). The structure of civil commitment: Patterns, pressures, and interactions in mental health legislation. *Law and Human Behavior*, 7, 1-18.
- Wexler, D. (1992). Putting mental health into mental health law: Therapeutic jurisprudence. *Law and Human Behavior*, 16, 27-38.
- Williamson, E. (2004). *Demographic Snapshot of State Trial Court Judges: 1979 and 2004*. American University. Retrieved on October 15, 2011 at http://www.naspaa.org/initiatives/paa/pdf/Erin_Williamson.pdf
- Winick, B. J. (2005). *Civil commitment. A therapeutic jurisprudence model*. Durham, NC: Carolina Academic Press.
- Wold, J.T. (1974). Political orientations, social backgrounds, and role perceptions of State Supreme Court judges. *The Western Political Quarterly*, 27, 239-248.
- Wrightsmann, L.S. (1999). *Judicial decision making. Is psychology relevant?* New York: Kluwer Academic/Plenum Publishers.
- Wyatt v. Stickney, 325 F. Supp. 781 (1971).
- Yamagishi, K. (1997). When a 12.86% mortality is more dangerous than 24.14%: Implications for risk communication. *Applied Cognitive Psychology*, 11, 495 – 506.

Appendix A

Strategies to Increase the Validity of Survey Research

One of the most efficient ways to increase the validity of survey research is through a mixed-mode design (i.e., using more than one way to gather data; Dillman, Smyth, & Christian, 2009). For this reason, participants could complete an electronic or paper version of the survey. Participants who completed the survey online did so through Survey Monkey (2007), a widely used software program designed to allow one to create customized online surveys. By using a mixed-mode design, this study ensured that judges who lack technological skills had the same opportunity to participate as judges with technological skills. Thus, this study recruited from both the subset of potential participants who prefer to complete tasks electronically and the subset of participants who prefer to complete the paper-version of tasks. Finally, a mixed-mode design was beneficial because response rates increase when potential participants are given different options of how to respond to a survey (Dillman et al., 2009).

The mixed-mode design decreases three of the four sources of errors that tend to plague survey research: (a) coverage, (b) sampling, and (c) nonresponse (Dillman et al., 2009). Coverage error occurs when not all members have a chance of being included in a survey sample. Sampling error refers to surveying only a subset of the entire population. This error exists as part of all sample surveys because not everyone in the true population is sampled (Dillman et al., 2009). Nonresponse errors occur when potential participants who respond to the survey are different from those who do not respond in a way that could impact the study. Mixed-mode design is unable to decrease the fourth source of error, which is measurement (i.e., when a participant's answer is inaccurate or imprecise). However, this study utilized various strategies designed to reduce the four potential sources of errors (see Table A1 below for description of these strategies). For example, the measures of this study were pilot tested, which decreases potential measurement error.

Table A1

Strategies to Decrease Potential Survey Errors

Purpose	Strategy	Rationale behind strategy
Decrease nonresponse error (i.e., when potential participants who respond to the survey are different from those who do not respond)	Mixed-mode survey design (i.e., electronic and paper-and-pencil surveys)	Mixed-mode survey design is the most important way to increase response rates because it gives individuals options in how to respond to a survey (Dillman et al., 2009).
	Minimized the time it takes to complete the survey by designing the survey to take less than 10 minutes to complete.	Potential participants are hesitant to complete surveys that require too much time to complete (Dillman et al., 2009).
	Judges received a follow-up email or letter reminding them about the study.	Research has indicated that the number of contacts influence response rates (Cook, Heath, & Thompson, 2000; Dillman et al., 2009; Sheehan & Hoy, 1999).
	Deadlines for participation were not stated.	Deadlines tend to decrease response rates for internet surveys especially among older participants (Goritz & Stieger, 2009).
	Recruitment emails were sent from a University of Alabama (UA) email account. Recruitment letters were sent in envelopes with the UA logo on the outside. The fact that study has been approved by UA IRB was mentioned in the recruitment email or letter.	Research has indicated sponsorship increases response rates (Dillman et al., 2009).
	This study used mail-merge software that allowed for the creation of personalized emails. Paper letters were personalized by using a Microsoft document template.	Research has indicated that personalized correspondence tends to increase response rates (Cook et al., 2000; Dillman et al., 2009).
	The recruitment emails and letters were designed to appeal to various types of potential participants by containing phrases that (a) asked for help, (b) showed positive regard, (c) said thank you, (d) supported group values, (e) emphasized saliency of topic, (f) avoided subordinate language, and (g) ensured confidentiality.	Research has indicated that potential participants differ with regard to which aspects of a survey motivate them to participate (Dillman et al., 2009; Groves, Singer, & Corning, 2000). For some potential participants, the saliency of the topic may be important but for others, the emotional satisfaction of responding to a researcher's plea for help may be more important.
Decrease sampling error (i.e., when only a subset of the population is surveyed)	The list of potential participants contained all 9,323 judges in the survey population	Dillman and colleagues (2009) indicated that if individuals in the survey population are left off eligible participants list, then the researcher would be gathering data from a subset of the population.

	Probability sampling equations were used to determine the appropriate sample size	These equations allow a researcher to generalize from the sample population to the true population with known precision (Dillman et al., 2009).
Decrease coverage error (i.e., when not all members have a chance of being included in the survey sample)	Mixed-mode survey design (i.e., electronic and paper-and-pencil surveys)	This strategy minimizes the coverage error attributable to lack of computer skills (Dillman et al., 2009).
	The list of potential participants contained all 9,323 judges in the survey population	Dillman and colleagues (2009) indicated that if individuals in the survey population are left off eligible participants list, then not every one would have the same opportunity to participate in the survey.
	The list were reviewed in order to ensure that no potential participant is on the list more than once	If some eligible participants were on the list more than once, then those participants would unintentionally have a higher likelihood of being in the study (Dillman et al., 2009).
	This study used mail-merge software which allowed for the creation of personalized emails.	This software allowed the emails to be sent individually instead of being sent as bulk emails. This technology decreased the likelihood firewall software blocked the email; or that the email was identified as and rerouted as junk mail.
Decrease measurement error (i.e., when a participant's answer is inaccurate or imprecise)	The wording of items and response options followed Dillman and colleagues' (2009) recommendations. Examples of the implementation of this strategy included: (1) allowing participants to skip questions 14 to 18 on demographics if they answered no to question 13, (2) using terms like 'probative' and 'respondent' (3) avoiding compound questions, (4) avoiding complex words, (5) avoiding general or abstract terms, (6) minimizing long questions, (7) using balanced likert scales, and (8) requiring that all likert scales have fewer than seven response options	Dillman and colleagues (2009) provide several recommendations that are suppose to make items and response options clear and concise for the participant. Recommendations included: (1) make sure the question applies to the participant, (2) make sure the sentence is technically accurate, (3) ask one question at a time, (4) use simple and familiar words, (5) use specific and concrete words, (6) use few words as possible, (7) avoid bias from unequal comparisons, and (8) limit scale length.
	A pilot-version of the study was conducted in order to refine items and instructions.	Dillman and colleagues (2009) recommended that surveys be pilot tested in order to gauge clarity of instructions and items.
	The measures contained items that have appeared in relevant literature.	Construct validity was established through use of items relevant to the construct (Dillman et al., 2009)

In order to increase response rate and decrease possible nonresponse error (i.e., when participants who respond are different from those who do not respond), this study incorporated the following empirically-supported recruitment strategies:

- 1) Dillman and colleagues (2009) indicated that using a mixed-mode survey design (i.e., electronic and paper-and-pencil surveys) is the most important way to increase response rates.
- 2) Research has indicated that the number of contacts influence response rates (Cook, Heath, & Thompson, 2000; Dillman et al., 2009; Sheehan & Hoy, 1999). As previously mentioned, judges received a follow-up email or letter reminding them about the study.
- 3) Deadlines for participation were stated because deadlines tend to decrease response rates for internet surveys, especially among older participants (Goritz & Stieger, 2009).
- 4) Research has indicated sponsorship increases response rates (Dillman et al., 2009). For this reason, the recruitment emails were sent from a University of Alabama (UA) email account and used a UA email template (see Appendices C and E). In addition, recruitment letters were printed on UA stationary, and the envelopes and follow-up postcards included UA logos (see Appendices D and F). The fact the study has been approved by UA Institutional Review Board (IRB) was mentioned in the recruitment email or letter and necessary contact information for IRB was provided.
- 5) Research has indicated that personalized correspondence tends to increase response rates (Cook et al., 2000; Dillman et al., 2009). For this reason, paper letters were personalized by using a Microsoft document template and emails were personalized by using the mail-merge feature of Microsoft Outlook. In addition to increasing response rates, the use of mail-merge software decreased possible coverage error. Specifically, this feature allowed

the emails to be sent individually, instead of being sent as bulk emails. This technology decreased the likelihood firewall software blocked the email; or that the email was identified as and rerouted as junk mail (Microsoft Corporation, 2010).

- 6) Research has indicated that potential participants differ with regard to which aspects of a survey motivate them to participate (Dillman et al., 2009; Groves, Singer, & Corning, 2000). For some potential participants, the saliency of the topic may be important but for others, the emotional satisfaction of responding to a researcher's request for help may be more important. For this reason, the recruitment emails and letters were designed to appeal to multiple personality characteristics by containing recommended phrases (see Figure A1 for illustration of this strategy).

Figure A1

Recruitment Wording Meant to Appeal to Multiple Personality Characteristics

Personalized correspondence increases response rates (Cook et al., 2000, Dillman et al., 2009) To JUDGES NAME Appeals to people's helping tendencies and pointing out the survey is national increases response rates (Groves et al., 1992; Mowen & Cialdini, 1980)

You are being asked to help with a brief national survey about judicial decision making in involuntary civil commitments (because you are a judge who has jurisdiction over civil commitment proceedings. This survey entitled *Community Violence Risk: Judicial Decision Making in Civil Commitments*, is being conducted by Stephanie Evans, a graduate student from The University of Alabama.

Pointing out why they have been chosen increases response rates (Dillman et al., 2009)

Using subordinate language and pointing out the survey is short increases response rates (Dillman et al., 2009)

Mentioning their group values/roles increases response rates (Dillman et al., 2009) Because you are a busy person, this survey was designed to take less than 10 minutes. The purpose of this study is to find out how risk communication messages and personal attributes impact judicial decision making. This survey contains questions about the probative value of risk communication messages, about hypothetical rulings, about your legal philosophy, and about your past experiences with civil commitments.

Your role as an essential member of the judicial system makes your input invaluable. Completing this survey will give the mental health community guidance on how clinicians can better serve you and your fellow judges.

Explaining how the results will benefit them increases response rates (Groves et al., 1992) 2,500 judges will be asked to participate in this survey. There will be no direct benefit to you but this survey has the potential to benefit all civil commitment judges. The risk of this study is that some of the questions may make you uncomfortable or you may find that a 10 minute survey gets in the way of you completing your work. In either case, you may discontinue the study at any time and your data will be discarded. Also, you may skip any question you do not want to answer. You will be provided a prepaid addressed envelope for your convenience, so there will be no cost to you.

Ensuring confidentiality increases response rates (Dillman et al., 2009)

It is important you, JUDGES NAME feel comfortable completing this survey. For this reason, this survey is completely confidential. Confidentiality will be ensured by not asking you to provide any identifying information. There will be no way to know whether you completed the survey or declined this request.

If you have questions about this study, please contact Stephanie Evans at 205-348-4014 (or via e-mail at saevars@crimson.ua.edu) or Dr. Karen Salekin (faculty advisor for the current study) at 205-348-0679 (or ksalekin@bama.ua.edu). If you have questions about your rights as a research participant, contact the University of Alabama Institutional Review Board at (205) 348-8461 or toll-free at 1-877-820-3066. If you are interested in completing the study online instead of through this paper-and-pencil version, you can complete the survey by going to the following website. *Appropriate survey monkey URL inserted here*

YOUR PARTICIPATION IS COMPLETELY VOLUNTARY. You are free to decline my request to participate or stop participating any time before you submit your answers.

If you understand the statements above, are at least 21 years old, and freely consent to be in this study, please complete the survey and return it in the prepaid addressed envelope provided.

Many thanks in advance for participation in this study. Saying thank you increases response rates (Dillman et al., 2009)

Kind regards,
Stephanie Evans, M.A.


THE UNIVERSITY OF ALABAMA
ARTS & SCIENCES


S/S Gordon Palmer
Box 870348
Tuscaloosa, Alabama 35487-0348
(205) 348-5083
Fax (205) 348-8648
www.ua.edu/psychology

Appendix B

Information about Pilot Study

A pilot version of this study was conducted in the Summer of 2010. The main purpose of the pilot study was to estimate response rates and to investigate the reliability of the legal philosophy, role orientations, and FFN scales. In addition, participants were provided an opportunity to offer suggestions or provide feedback regarding the clarity of instructions and items.

The sample population for this study were judges who the NCSC (2009) identified as having jurisdiction over involuntary civil commitments. Prior to the recruitment phase of the pilot study, contact information for judges was obtained by: (a) internet searches via publically accessible websites and (b) searches of directories¹³ located at the University of Alabama Law Library. Mailing addresses were obtained for all 12,821 judges deemed eligible for this study. In addition, email addresses were obtained for 3,722 of the 12,821 judges (i.e., 29% had available email addresses).

Two hundred potential participants were randomly selected from the list of 12,821 eligible judges. Recruitment emails were sent to potential participants with available email addresses (of the 200 judges randomly selected, 60 judges had email addresses). Five of these recruitment emails were returned as undeliverable. Thus, paper recruitment letters were sent to 145 potential participants (i.e., 140 judges without email addresses plus 5 judges with incorrect email addresses). Two recruitment letters were returned as undeliverable for unknown reasons. The recruitment emails and letters were sent between June 14 and June 18, 2010. Two weeks later (i.e., between June 29 and July 2, 2010) a follow-up email or postcard was sent to remind judges about the study or thank them for their participation. The recruitment emails and letters were sent on different days and times to investigate whether these factors had an impact on response rates (see Tables A2 and A3).

¹³ Specifically, the following directories were used *The American Bench: Judges of the Nation* (2010), *Directory of Minority Judges in the United States* (2008), and *BNA's Directory of State and Federal Courts, Judges, and Clerk* (2010)

Table A2

Response Rates for Recruitment Letters

Contact Sent	Surveys completed during initial phase	Surveys completed during follow-up phase	Clarification responses	Undeliverable letters
Monday	2	0	0	0
Tuesday	6	0	0	0
Wednesday	2	0	1	1
Thursday	0	0	1	1
Friday	3	2	0	0

Twenty-two judges responded to the recruitment letters/emails by completing the study (11% response rate). Specifically, 15 judges responded to the recruitment letters (10.4% response rate) and 7 judges responded to the recruitment emails (12.72%). Of the 22 judges who responded, 17 responded during the initial contact phase and 5 responded during the follow-up phase. The follow-up phase increased the paper-version response rate by 13% (2 follow-up response divided by total of 15) and the electronic responses by 43% (3 follow-up response divided by total of 7). It should be noted, however, that the follow-up postcards included the electronic address of the survey. As such, it is possible some of the electronic responses were from judges who originally received the paper version of the study. A review of the response rates indicated judges were more likely to complete the survey when the recruitment letter was sent on Tuesday or Friday (see Table A2 for responses to recruitment letter). Judges were more likely to complete the survey when the recruitment email was sent on Thursday between 8:15 and 8:45 AM CDT (see Table A3 for responses to recruitment email).

Table A3

Response Rates for Recruitment Emails

Contact Sent	Surveys completed during initial phase	Surveys completed during follow-up phase	Clarification responses
Monday	2	0	0
Tuesday	0	0	2
Wednesday	0	2	0
Thursday	2	1	0
Friday	0	0	0

In addition to the 22 judges who completed the study, four judges informed the researcher they would like to participate, but they did not oversee civil commitment proceedings. Three of these judges provided a thorough explanation about which judicial departments or divisions oversaw civil commitment proceedings in their state. For example, the NCSC (2009) indicated that in Pennsylvania the Court of Common Pleas has jurisdiction over civil commitments. Feedback from one judge indicated that in some districts in Pennsylvania, only judges in the Orphan’s Court Division of the Court of Common Pleas oversee civil commitments.

Although the construct validity of the scales was going to be investigated via factor analyses, statistical manuals (DeCoster & Claypool, 2004; Pallant, 2007; Tabachnick & Fidell, 2007) indicated factor analyses with sample sizes of less than 100 subjects would produce unstable and potentially erroneous results. For this reason, internal consistency was the main analysis used to investigate the extent to which the items on the scale measured the same construct. Finally, skewness, kurtosis, mean, and response range for the scales were gathered to identify any potential restrictions (or abnormalities) in the distributions of scores.

The Legal Philosophy scale (see Table A4) demonstrated inadequate internal consistency ($\alpha = .222$) when all five items were included. Cronbach alpha coefficients of over .7 are indicative of adequate internal consistency and alphas coefficients of .5 are considered common in scales with less

than five items (DeCoster & Claypool, 2004; Pallant, 2007). Excluding two items from the scale¹⁴ increased the internal consistency to a Cronbach alpha coefficient of .7. The distribution of scores for this three-item scale was normal with a skew of 1.28 and a kurtosis of -.58.¹⁵ The pilot sample produced a mean of 10.38 ($SD = 4.8$) and total scores for the three-item version of the scale ranged from 4 to 21 (possible scores of 3 to 21).

Table A4

Legal Philosophy Scale Used in Pilot Study and Item Analysis

Original Legal Philosophy scale items*	Cronbach Alpha if Item is deleted
The purpose of civil commitments is to care for individuals who cannot care for themselves.	.277
The purpose of civil commitments is to protect the community from potentially dangerous individuals.**	.418
Being civilly committed to an inpatient psychiatric facility is punitive for most mentally ill respondents.	-.098
Being civilly committed to an inpatient psychiatric facility is beneficial for most mentally ill respondents.**	.542
Civil commitments are more about helping individuals who cannot care for themselves than protecting the public from potentially dangerous individuals.	-.169

* Participants rate their agreement or disagreement of each item through a seven-point Likert scale (1 = *Strongly disagree* to 7 = *Strongly agree*).

**These items are excluded from the subsequent version of the scale.

¹⁴The two items excluded were: “Being civilly committed to an inpatient psychiatric facility is beneficial for most mentally ill respondents” and “The purpose of civil commitments is to protect the community from potentially dangerous individuals.” Item analysis indicated that the removal of these two items would significantly increase the internal consistency of the scale.

¹⁵ Skew within the -2 to 2 range and kurtosis within -2 to 2 suggests the data is normally distributed (Garson, 2010).

Table A5

Role Orientation Scale Used in Pilot Study and Factor Analysis

Original Role Orientation Items*	Rotated Component Matrix ^a	
	Role Orientation	Legal Philosophy
A good judge is one who sticks as closely to precedents as possible.**	.083 (weak)	.489 (moderate)
Judges should be allowed great discretion in decision making to ensure that their decisions are just.	.809 (large)	-.340 (weak)
It is wrong for a judge to allow his personal philosophy to influence his decisions.	.744 (large)	.296 (weak)
Precedents and statutes are only a few of the factors which should influence judges' decisions.	.656 (large)	.210 (weak)
It is better to strictly interpret statutes than allow discretion in decision making.	.837 (large)	.320 (moderate)

* Participants rate their agreement or disagreement of each item through a seven-point Likert scale (1 = *Strongly disagree* to 7 = *Strongly agree*).

**This item is excluded from the subsequent version of the scale.

^a Note: An exploratory factor analysis was conducted in which 2 factors were extracted from the items. Although not presented in the table, the three items from the legal philosophy scale were included in the factor analysis. The results of the factor analysis are presented via a Rotated Component Matrix, which indicates the extent to which each item loads on the two factors. Factor loadings less than .3 are considered weak, loadings between .3 and .6 are considered moderate, and loadings greater than .6 are considered to be large (DeCoster & Claypool, 2004).

The five-item Role Orientation scale (see Table A5) demonstrated good internal consistency with a Cronbach alpha coefficient of .71. To decrease the correlation between the Role Orientation scale and the three-item version of the Legal Philosophy scale, one item was excluded from the Role Orientation scale.¹⁶ Although neither correlation was significant¹⁷, this modification decreased the

¹⁶ The excluded item was "A good judge is one who sticks as closely to precedents as possible." This item was removed because results from an exploratory factorial analysis indicated that this item was more related to legal philosophy than role orientation. Although factorial analysis results should not be reported for sample sizes of less than 100 subjects, informal correspondence with statistician Jamie DeCoster indicated factor analysis is appropriate to use for exploratory purposes (J. DeCoster, personal communication, July 14, 2010).

¹⁷ The correlation ($r = .38$) between the five-item version of the Role Orientation Scale and the three-item version of the Legal Philosophy Scale was nonsignificant at $p = .09$. The correlation between the three-item version of the Role

variance shared between the scales from 14.4% to 10.2%. The distribution of the scores for the four-item version of the scale was normal with a skew of $-.36$ and a kurtosis of $-.66$. The Cronbach alpha coefficient was $.75$ for the four-item version of this scale. The pilot sample produced a mean of 15.57 ($SD = 5.81$) and the total scores for this scale ranged from 5 to 25 (possible scores of 4 to 28).

The FFN scale (see Table A6), which originally contained ten items, demonstrated an internal consistency of $.51$. Because Cronbach alpha coefficients should be above $.6$ in order for a scale to have “acceptable” internal consistency, three items¹⁸ were excluded. The seven-item version of the scale produced a Cronbach alpha coefficient of $.69$ and a normal distribution of scores (skew = 1.61 and kurtosis = $.76$). The pilot sample produced a mean of 26.67 ($SD = 7.03$) and total scores for the seven-item version of this scale ranged from 17 to 44 (possible scores of 7 to 49).

Orientation Scale and the four-item version of the Legal Philosophy Scale was lower ($r = .32, p = .15$).

¹⁸ The three items were: “I tend to err on the side of caution and commit respondents who present only a small chance of being violent,” “Whenever I make civil commitment decisions, I find similarities between the present respondent and prior respondents who ended up committing violent acts,” and “Community members would judge me harshly if I declined commitment for a respondent who was subsequently violent.” Item analysis indicated the removal of these items would significantly increase the internal consistency of the scale.

Table A6

Fear of False Negatives Scale Used in Pilot Study and Item Analysis

Original FFN scale items	Cronbach Alpha if Item is deleted
If I declined commitment for a respondent who was subsequently violent, the media and community members would find out that I had been the one who sent the respondent back into the community.	.530
I tend to err on the side of caution and commit respondents who present only a small chance of being violent.**	.573
The worst consequence of a civil commitment hearing is to decline commitment for a respondent who shortly afterwards committed a violent act.	.452
The worst consequence of a civil commitment hearing is to commit a respondent who would have never actually committed a violent act.	.488
Whenever I make civil commitment decisions, I find similarities between the present respondent and prior respondents who ended up committing violent acts. **	.560
My career as a judge would be terminated if I declined commitment for a respondent who was subsequently violent.	.448
Community members would judge me harshly if I declined commitment for a respondent who was subsequently violent.**	.552
If I declined commitment for a respondent who was subsequently violent, the image and case characteristics of the respondent would be burned in my mind forever.	.365
I always consider the possibility that if I deny an involuntary civil commitment the respondent may end up committing a violent act in the community.	.443
If I denied involuntary civil commitment in a case and shortly afterwards, the respondent committed a violent act, I would feel guilty and remorseful about my decision.	.380

* Participants rate their agreement or disagreement of each item through a seven-point Likert scale (1 = *Strongly disagree* to 7 = *Strongly agree*).

**These items are excluded from the subsequent version of the scale.

In addition to estimating response rates and investigating the construct validity of the scales, a purpose of this pilot study was to gather feedback or suggestions from the participants. Eight out of the twenty-two participants who completed the survey provided feedback about the study or clarification about themselves. Two out of the eight provided clarification of their role in civil commitment proceedings (i.e., one stated he oversaw mostly sexually violent predator commitments in which the jury made the determination, and the other participant stated he makes judicial determinations based on mental health commissioner reports). One participant pointed out judges rarely know what happens to respondents after they leave the courtroom, while another participant opined that no judge would allow fear of public scrutiny to impact his or her decision. The remaining four participants expressed that they desired clarification or additional information in the risk vignettes. Specifically, two participants wanted clarification or more information about whether the respondent posed a danger to himself. One wanted clarification on the term “need for treatment,” while another requested clarification of to whom the respondent had been violent toward in the past (i.e., family members or strangers).

Based on the information gathered from this pilot study, the following considerations were taken into account when designing the current study:

- 1) Decisions about how many judges to contact for the current study were based on the assumption that a response rate of at least 11% would be attained.
- 2) Based on the pattern of response rates in the pilot study, recruitment letters for the current study were sent on Tuesdays or Fridays. Recruitment emails were sent on Thursdays around 8:30 AM CDT.

- 3) The Legal Philosophy scale included three items because this version of the scale demonstrated adequate internal consistency and a normal distribution during the pilot study.
- 4) The Role Orientation scale included four items because this version of the scale was less correlated with the construct legal philosophy. In addition, this scale demonstrated adequate internal consistency and a normal distribution during the pilot study.
- 5) The FFN scale included seven items because this version of the scale demonstrated acceptable internal consistency and a normal distribution during the pilot study.
- 6) The wording of the risk vignettes was modified in order to address participants' concerns.
- 7) Clarification provided during the recruitment phase indicated there are probably less than 12,821 judges who oversee civil commitment proceedings in practice. As a result, the list of eligible participants was reduced in order to exclude judges who do not actually oversee civil commitments.

Appendix C

Recruitment Email

(“Research Invitation” on the message line of an e-mail)

THE UNIVERSITY OF ALABAMA

To JUDGES NAME:

You are being asked to help with a brief national survey about judicial decision making in involuntary civil commitments because you are a judge who has jurisdiction over civil commitment proceedings. This survey entitled *Communicating Violence Risk: Judicial Decision Making in Civil Commitments*, is being conducted by Stephanie Evans, a graduate student from The University of Alabama.

Because you are a busy person, this survey was designed to take **less than 10 minutes**. The purpose of this study is to find out how risk communication messages and personal attributes impact judicial decision making. This survey contains questions about the probative value of risk communication messages, about hypothetical rulings, about your legal philosophy, and about your past experiences with civil commitments.

Your role as an essential member of the judicial system makes your input invaluable. Completing this survey will give the mental health community **guidance on how clinicians can better serve you and your fellow judges**. 2,500 judges will be asked to participate in this survey. There will be no direct benefit to you but this survey has the potential to benefit all civil commitment judges. The risk of this study is that some of the questions may make you uncomfortable or you may find that a 10 minute survey gets in the way of you completing your work. In either case, you may discontinue the study at any time and your data will be discarded. Also, you may skip any question you do not want to answer.

It is important you, JUDGES NAME feel comfortable completing this survey. For this reason, this survey is **completely confidential**. Confidentiality will be ensured by not asking you to provide any identifying information and using SSL encryption to protect your responses. There will be no way to know whether you completed the survey or declined this request.

If you have questions about this study, please contact Stephanie Evans at 205-348-4014 (or via e-mail at saevans@crimson.ua.edu) or Dr. Karen Salekin (faculty advisor for the current study) at 205-348-0679 (or ksalekin@bama.ua.edu). If you have questions about your rights as a research participant, contact the University of Alabama Institutional Review Board at (205) 348-8461 or toll-free at 1-877-820-3066. **If you are interested in completing the study but would prefer a paper-and-pencil version, please let us know by emailing** saevans@crimson.ua.edu.

YOUR PARTICIPATION IS COMPLETELY VOLUNTARY. You are free to decline my request to participate or stop participating any time before you submit your answers.

If you understand the statements above, are at least 21 years old, and freely consent to be in this study, click on the following link button to begin.

Appropriate survey monkey URL inserted here

Many thanks in advance for participation in this study.

Kind regards,
Stephanie Evans, M.A.

Department of Psychology
Box 870348
Tuscaloosa, AL 35487-0348

Appendix D

Recruitment Letter

College of Arts and Science

Department of Psychology



DATE

To JUDGES NAME:

You are being asked to help with a brief national survey about judicial decision making in involuntary civil commitments because you are a judge who has jurisdiction over civil commitment proceedings. This survey entitled *Communicating Violence Risk: Judicial Decision Making in Civil Commitments*, is being conducted by Stephanie Evans, a graduate student from The University of Alabama.

Because you are a busy person, this survey was designed to **take less than 10 minutes**. The purpose of this study is to find out how risk communication messages and personal attributes impact judicial decision making. This survey contains questions about the probative value of risk communication messages, about hypothetical rulings, about your legal philosophy, and about your past experiences with civil commitments.

Your role as an essential member of the judicial system makes your input invaluable. Completing this survey will give the mental health community **guidance on how clinicians can better serve you and your fellow judges**. 2,500 judges will be asked to participate in this survey. There will be no direct benefit to you but this survey has the potential to benefit all civil commitment judges. The risk of this study is that some of the questions may make you uncomfortable or you may find that a 10 minute survey gets in the way of you completing your work. In either case, you may discontinue the study at any time and your data will be discarded. Also, you may skip any question you do not want to answer. You will be **provided a prepaid addressed envelope** for your convenience, so there will be no cost to you.

It is important you, JUDGES NAME feel comfortable completing this survey. For this reason, this survey is **completely confidential**. Confidentiality will be **ensured** by not asking you to provide any identifying information. There will be no way to know whether you completed the survey or declined this request.

If you have questions about this study, please contact Stephanie Evans at 205-348-4014 (or via e-mail at saevans@crimson.ua.edu) or Dr. Karen Salekin (faculty advisor for the current study) at 205-348-0679 (or ksalekin@bama.ua.edu). If you have questions about your rights as a research participant, contact the University of Alabama Institutional Review Board at (205) 348-9461 or toll-free at 1-877-820-3066. If you are interested in completing the study online instead of through this paper-and-pencil version, you can complete the survey by going to the following website. *Appropriate survey monkey URL inserted here*



318 Gordon Palmer
Box 870318

Tuscaloosa, Alabama 35487-0318
(205) 348-5085

Fax (205) 348-8648
www.ua.edu/psychology

YOUR PARTICIPATION IS COMPLETELY VOLUNTARY. You are free to decline my request to participate or stop participating any time before you submit your answers.

If you understand the statements above, are at least 21 years old, and freely consent to be in this study, please complete the survey and return it in the prepaid addressed envelope provided.

Many thanks in advance for participation in this study.

Kind regards,
Stephanie Evans, M.A.

Appendix E

Follow-up Email

THE UNIVERSITY OF ALABAMA

To Hon. LASTNAME:

Two weeks ago an email was sent to you asking your help with a brief survey about judicial decision making in involuntary civil commitments.

If you completed the survey, thank you so much for your assistance in this study. Your input is invaluable. If you have not completed the survey, there is still time to participate. You can complete the survey online by going to the following website: *Appropriate surveymonkey.com link inserted here*. Or you can email saevans@bama.ua.edu and request a paper and pencil version of the survey.

Thank you,

Stephanie Evans, M.A.

Department of Psychology
Box 870348
Tuscaloosa, AL 35487-0348

Appendix F

Follow-up Postcard

UNIVERSITY OF ALABAMA
DEPARTMENT OF PSYCHOLOGY
C/O STEPHANIE EVANS
BOX 870348
TUSCALOOSA AL 35487-0348



JUDGE'S NAME
FIRST ADDRESS LINE
SECOND ADDRESS LINE
CITY STATE ZIP CODE

To JUDGE NAME,

Two weeks ago a letter was sent to you asking your help with a brief survey about judicial decision making in involuntary civil commitments.

If you completed the survey, thank you so much for your assistance in this study. Your input is invaluable. If you have not completed the survey, there is still time to participate. You complete and return the survey we previously provided. Or you can email saevans@crimson.ua.edu and request a new survey be sent to you. Additionally, you can complete the survey online by going to the following website: *Appropriate surveymonkey.com link inserted here.*

Thank you,
Stephanie Evans

Appendix G

Demographics Questionnaire

1. What is your gender?
 - a. Male
 - b. Female
2. What is your age?
_____ years old
3. What is your ethnic identity/race?
 - a. Caucasian
 - b. African-American
 - c. Latino/Latina
 - d. Asian
 - e. Other _____
4. What is the highest level of education you completed?
 - a. High School/GED or less
 - b. Some college
 - c. Bachelor Degree
 - d. Graduate Degree
 - e. Juris Doctor (JD)
5. Are you an active judge or are you a retired judge?
 - a. Active judge
 - b. Retired judge
6. How many years have you been (or were) a judge?
_____ years
7. How many years have you been (or were) a judge who has jurisdiction over involuntary civil commitments?
_____ years
8. During you entire career, how many civil commitment cases do you estimate you have presided over?
_____ cases
9. During you entire career, has there ever been a case in which you denied a civil commitment and you know that the respondent subsequently committed a violent act?
 - a. Yes
If you circled yes, in how many cases did this happen?
_____ cases
 - b. No
10. In what state are you currently employed as a judge? *Or if you are a retired judge, in what state were you most recently employed as a judge?*

11. *States differ in which type of judges or courts have legal jurisdiction over involuntary civil commitments. What type of judge are/were you or what court are/were you affiliated with?*
 - a. Magistrate
 - b. Probate
 - c. Superior
 - d. District
 - e. Circuit
 - f. Court of chancery
 - g. Other _____
12. Were you elected by the public or appointed to your most recent position of judge?
 - a. Appointed
 - b. Elected
 - c. Other _____
13. Prior to being a judge, were you an attorney?
 - a. Yes
 - b. No

Appendix H

Risk Vignettes

Risk description model

Introduction to all vignettes: You are the judge presiding over the civil commitment hearing of Mr. Doe/Smith who has been assessed for the risk of future violence by a clinician. Last week, Mr. Doe's/Smith's family filed a petition for civil commitment because he began to respond to auditory hallucinations and express some paranoid thoughts. After listening to the testimony of family members, you are positive that the following criteria have been sufficiently established: (a) the person has a bonafide mental illness and (b) the person lacks the capacity to make treatment decisions. Now you must decide whether the "danger to others" criterion is met based on the following information provided by the clinician who conducted the risk assessment.

Condition One:

High risk

Mr. Smith is a 24-year-old Caucasian male who has been diagnosed, for several years, with Schizophrenia. His family reported he has not been taking his psychiatric medication for over a month. He has been in and out of substance abuse treatment since he was 15 years old and he continues to heavily use cocaine and alcohol on a daily basis. He has a history of violent behavior toward family members and strangers. With regard to protective factors, Mr. Smith's family is considered to be supportive.

Condition Six:

Low risk

Mr. Doe is a 35-year-old Caucasian male who has been diagnosed, for several years, with Schizophrenia. His family reported he has not been taking his psychiatric medication for over a month. He has no history of substance abuse. He has no history of violent behavior toward family members or strangers. With regard to protective factors, Mr. Doe's family is considered to be supportive.

Risk prediction model with categorical format

Introduction to all vignettes: You are the judge presiding over the civil commitment hearing of Mr. Doe/Smith who has been assessed for the risk of future violence by a clinician. Last week, Mr. Doe's/Smith's family filed a petition for civil commitment because he began to respond to auditory hallucinations and express some paranoid thoughts. After listening to the testimony of family members, you are positive that the following criteria have been sufficiently established: (a) the person has a bonafide mental illness and (b) the person lacks the capacity to make treatment decisions. Now you must decide whether the "danger to others" criterion is met based on the following information provided by the clinician who conducted the risk assessment.

Condition Two:

High risk

Mr. Smith is a 24-year-old Caucasian male who has been diagnosed, for several years, with Schizophrenia. His family reported he has not been taking his psychiatric medication for over a month. He has been in and out of substance abuse treatment since he was 15 years old and he continues to heavily use cocaine and alcohol on a daily basis. He has a history of violent behavior toward family members and strangers. With regard to protective factors, Mr. Smith's family is considered to be supportive. Considering the risk and protective factors previously described, Mr. Smith presents a high risk of violence in the near future.

Condition Seven:

Low risk

Mr. Doe is a 35-year-old Caucasian male who has been diagnosed, for several years, with Schizophrenia. His family reported he has not been taking his psychiatric medication for over a month. He has no history of substance abuse. He has no history of violent behavior toward family members or strangers. With regard to protective factors, Mr. Doe's family is considered to be supportive. Considering the risk and protective factors previously described, Mr. Doe presents a low risk of violence in the near future.

Risk prediction model with frequency format

Introduction to all vignettes: You are the judge presiding over the civil commitment hearing of Mr. Doe/Smith who has been assessed for the risk of future violence by a clinician. Last week, Mr. Doe's/Smith's family filed a petition for civil commitment because he began to respond to auditory hallucinations and express some paranoid thoughts. After listening to the testimony of family members, you are positive that the following criteria have been sufficiently established: (a) the person has a bonafide mental illness and (b) the person lacks the capacity to make treatment decisions. Now you must decide whether the "danger to others" criterion is met based on the following information provided by the clinician who conducted the risk assessment.

Condition Three:

High risk

Mr. Smith is a 24-year-old Caucasian male who has been diagnosed, for several years, with Schizophrenia. His family reported he has not been taking his psychiatric medication for over a month. He has been in and out of substance abuse treatment since he was 15 years old and he continues to heavily use cocaine and alcohol on a daily basis. He has a history of violent behavior toward family members and strangers. With regard to protective factors, Mr. Smith's family is considered to be supportive. Considering the risk and protective factors previously described, 76 out of every 100 people with similar features as Mr. Smith will commit a violent act toward others in the near future.

Condition Eight:

Low risk

Mr. Doe is a 35-year-old Caucasian male who has been diagnosed, for several years, with Schizophrenia. His family reported he has not been taking his psychiatric medication for over a month. He has no history of substance abuse. He has no history of violent behavior toward family members or strangers. With regard to protective factors, Mr. Doe's family is considered to be supportive. Considering the risk and protective factors previously described, 29 out of every 100 people with similar features as Mr. Doe will commit a violent act toward others in the near future.

Risk prediction model with probabilistic format

Introduction to all vignettes: You are the judge presiding over the civil commitment hearing of Mr. Doe/Smith who has been assessed for the risk of future violence by a clinician. Last week, Mr. Doe's/Smith's family filed a petition for civil commitment because he began to respond to auditory hallucinations and express some paranoid thoughts. After listening to the testimony of family members, you are positive that the following criteria have been sufficiently established: (a) the person has a bonafide mental illness and (b) the person lacks the capacity to make treatment decisions. Now you must decide whether the "danger to others" criterion is met based on the following information provided by the clinician who conducted the risk assessment.

Condition Four:

High risk

Mr. Smith is a 24-year-old Caucasian male who has been diagnosed, for several years, with Schizophrenia. His family reported he has not been taking his psychiatric medication for over a month. He has been in and out of substance abuse treatment since he was 15 years old and he continues to heavily use cocaine and alcohol on a daily basis. He has a history of violent behavior toward family members and strangers. With regard to protective factors, Mr. Smith's family is considered to be supportive. Considering the risk and protective factors previously described, Mr. Smith is 76% likely to commit a violent act toward others in the near future.

Condition Nine:

Low risk

Mr. Doe is a 35-year-old Caucasian male who has been diagnosed, for several years, with Schizophrenia. His family reported he has not been taking his psychiatric medication for over a month. He has no history of substance abuse. He has no history of violent behavior toward family members or strangers. With regard to protective factors, Mr. Doe's family is considered to be supportive. Considering the risk and protective factors previously described, Mr. Doe is 29% likely to commit a violent act toward others in the near future.

Risk management model

Introduction to all vignettes: You are the judge presiding over the civil commitment hearing of Mr. Doe/Smith who has been assessed for the risk of future violence by a clinician. Last week, Mr. Doe's/Smith's family filed a petition for civil commitment because he began to respond to auditory hallucinations and express some paranoid thoughts. After listening to the testimony of family members, you are positive that the following criteria have been sufficiently established: (a) the person has a bonafide mental illness and (b) the person lacks the capacity to make treatment decisions. Now you must decide whether the "danger to others" criterion is met based on the following information provided by the clinician who conducted the risk assessment.

Condition Five:

High risk

Mr. Smith is a 24-year-old Caucasian male who has been diagnosed, for several years, with Schizophrenia. His family reported he has not been taking his psychiatric medication for over a month. He has been in and out of substance abuse treatment since he was 15 years old and he continues to heavily use cocaine and alcohol on a daily basis. He has a history of violent behavior toward family members and strangers. With regard to protective factors, Mr. Smith's family is considered to be supportive. Medication noncompliance and substance abuse tend to elevate an individual's risk to commit a violent act. To reduce potential risk for violence, Mr. Smith should receive long-acting injectable medication. To address his substance abuse, he should receive random drug testing and support from Alcoholics Anonymous (AA) or Narcotics Anonymous. Support from his family and possible employment placement should also be encouraged to order to decrease his substance abuse.

Condition Ten:

Low risk

Mr. Doe is a 35-year-old Caucasian male who has been diagnosed, for several years, with Schizophrenia. His family reported he has not been taking his psychiatric medication for over a month. He has no history of substance abuse. He has no history of violent behavior toward family members or strangers. With regard to protective factors, Mr. Doe's family is considered to be supportive. Medication noncompliance tends to elevate an individual's risk to commit a violent act. To reduce potential risk for violence, Mr. Doe should receive long-acting injectable medication.

Appendix I

Follow-up Questions to Risk Vignettes

1. To what extent do you agree or disagree that the “danger to others” criterion has been sufficiently met in this case?

Strongly Disagree	Somewhat Disagree	Slightly Disagree	Neither agree nor disagree	Slightly Agree	Somewhat Agree	Strongly Agree
○	○	○	○	○	○	○

2. Which of the responses best represents your ruling in this case? *Assume that the following criteria have been sufficiently established: (1) the person has a bona fide mental illness and (2) the person lacks the capacity to make treatment decisions.*

- a. The respondent should **not** be committed
- b. I am **unsure** whether the respondent should be committed to an **outpatient facility**, but I am sure that he should not be committed to an inpatient facility.
- c. The respondent should be committed to an **outpatient facility**.
- d. I am **unsure** whether the respondent should be committed to an **inpatient facility**, but I am sure that he meets the criteria for commitment to an outpatient facility.
- e. The respondent should be committed to an **inpatient facility**.

3. How probative do you believe the risk vignette was in helping you make your decisions about this case?

Not at all valuable	Slightly valuable	Somewhat valuable	Very valuable	Extremely valuable
○	○	○	○	○

Appendix J

Judicial Attribute Questionnaire (Role Orientation, Legal Philosophy, and Fear of False Negatives)

Please read the following statements and determine the extent to which you agree or disagree with the statement.

- If you **strongly disagree** with the statement
- If you **moderately disagree** with the statement
- If you **slightly disagree** with the statement
- If you are exactly and precisely **neutral** about the statement
- If you **slightly agree** with the statement
- If you **moderately agree** with the statement
- If you **strongly agree** with the statement

You may find that you may have different reactions to different parts of the statement. For example, you might very strongly agree with one idea in a statement but slightly agree with another idea in the same item. When this happens, please combine your reactions and shade or check the circle representing how you feel "on average".

		Strongly Disagree	Moderately Disagree	Slightly Disagree	Neutral	Slightly Agree	Moderately Agree	Strongly Agree
1.	Judges should be allowed great discretion in decision making to ensure that their decisions are just.	<input type="radio"/>						
2.	The worst consequence of a civil commitment hearing is to decline commitment for a respondent who shortly afterward committed a violent act.	<input type="radio"/>						
3.	It is wrong for a judge to allow his personal philosophy to influence his decisions.	<input type="radio"/>						
4.	It is better to strictly interpret statutes than allow discretion in decision making.	<input type="radio"/>						
5.	Being civilly committed to an inpatient psychiatric facility is punitive for most mentally ill respondents.	<input type="radio"/>						
6.	If I declined commitment for a respondent who was subsequently violent, the media and community members would find out that I had been the one who sent the respondent back into the community.	<input type="radio"/>						
7.	The purpose of civil commitments is to care for individuals who cannot care for themselves.	<input type="radio"/>						
8.	The worst consequence of a civil commitment hearing is to commit a respondent who would have never actually committed a violent act.	<input type="radio"/>						
9.	Civil commitments are more about helping individuals who cannot care for themselves than protecting the public from potentially dangerous individuals.	<input type="radio"/>						
10.	If I declined commitment for a respondent who was subsequently violent, the image and case characteristics of the respondent would be burned in my mind forever.	<input type="radio"/>						
11.	Precedents and statutes are only a few of the factors which should influence judges' decisions.	<input type="radio"/>						
12.	My career as a judge would be terminated if I declined commitment for a respondent who was subsequently violent.	<input type="radio"/>						
13.	I always consider the possibility that if I deny an involuntary civil commitment the respondent may end up committing a violent act in the community.	<input type="radio"/>						
14.	If I denied involuntary civil commitment in a case and shortly afterward, the respondent committed a violent act, I would feel guilty and remorseful about my decision.	<input type="radio"/>						

Appendix K

Additional Questions for Subsequent Study

- 1) What categorical terms would you prefer clinicians use when describing risk?
 - a. Low risk, moderate risk, or high risk
 - b. Low risk, average risk, or high risk
 - c. Low risk, moderate risk, high risk, or very high risk
 - d. Other: _____

- 2) The term “low risk” indicates that individual has a _____ percent chance of committing a violent behavior in the near future.

- 3) The term “moderate risk” indicates that individual has a _____ percent chance of committing a violent behavior in the near future.

- 4) The term “high risk” indicates that individual has a _____ percent chance of committing a violent behavior in the near future.

- 5) What is the lowest likelihood of future violent behavior that is sufficient to warrant involuntary civil commitment?
 - a. In terms of categories (Please circle your response):
 - i. Low risk
 - ii. Moderate risk
 - iii. High risk

 - b. In terms of probability (Please write in your response and use percentages): _____

Appendix L

Manipulation Check

A manipulation check was conducted in order to verify the construct validity of risk level in the vignettes. Twenty seven individuals (i.e., graduate students and forensic psychologists) were asked to participate in this manipulation check. Participants were asked to read a risk vignette and then rate Mr. Smith's/Doe's risk to commit a violent act in the near future via a seven-point response scale (1 = Low risk to 7 = High risk). Thirteen individuals were asked to rate the high risk vignette (i.e., Mr. Smith) and fourteen individuals were asked to rate the low risk vignette (i.e., Mr. Doe). Ten individuals rated the high risk vignette and obtained a mean of 5.4 (Mdn = 5.5, Modes = 5 and 6) and a standard deviation of 1.07. Thirteen individuals rated the low risk vignette and obtained a mean of 3.23 (Mdn = 3, Mode = 3) and a standard deviation of 1.59. Mr. Doe in the low risk vignette was rated as presenting a significantly lower risk than Mr. Smith in the high risk vignette, $t [21] = 3.70, p = .001$.

Appendix M

Judicial Attribute Questionnaire (Role Orientation, Legal Philosophy, and Fear of False Negatives)

Please read the following statements and determine the extent to which you agree or disagree with the statement.

- If you **strongly disagree** with the statement
- If you **moderately disagree** with the statement
- If you **slightly disagree** with the statement
- If you are exactly and precisely **neutral** about the statement
- If you **slightly agree** with the statement
- If you **moderately agree** with the statement
- If you **strongly agree** with the statement

You may find that you may have different reactions to different parts of the statement. For example, you might very strongly agree with one idea in a statement but slightly agree with another idea in the same item. When this happens, please combine your reactions and shade or check the circle representing how you feel "on average".

		Strongly Disagree	Moderately Disagree	Slightly Disagree	Neutral	Slightly Agree	Moderately Agree	Strongly Agree
Role Orientation	1. Judges should be allowed great discretion in decision making to ensure that their decisions are just.	○	○	○	○	○	○	○
	2. It is wrong for a judge to allow his personal philosophy to influence his decisions.	○	○	○	○	○	○	○
	3. Precedents and statutes are only a few of the factors which should influence judges' decisions.	○	○	○	○	○	○	○
	4. It is better to strictly interpret statutes than allow discretion in decision making.	○	○	○	○	○	○	○
Legal Philosophy	5. The purpose of civil commitments is to care for individuals who cannot care for themselves.	○	○	○	○	○	○	○
	6. Being civilly committed to an inpatient psychiatric facility is punitive for most mentally ill respondents.	○	○	○	○	○	○	○
	7. Civil commitments are more about helping individuals who cannot care for themselves than protecting the public from potentially dangerous individuals.	○	○	○	○	○	○	○
	8. If I declined commitment for a respondent who was subsequently violent, the media and community members would find out that I had been the one who sent the respondent back into the community.	○	○	○	○	○	○	○
FFN	9. The worst consequence of a civil commitment hearing is to decline commitment for a respondent who shortly afterward committed a violent act.	○	○	○	○	○	○	○
	10. The worst consequence of a civil commitment hearing is to commit a respondent who would have never actually committed a violent act.	○	○	○	○	○	○	○
	11. My career as a judge would be terminated if I declined commitment for a respondent who was subsequently violent.	○	○	○	○	○	○	○
	12. If I declined commitment for a respondent who was subsequently violent, the image and case characteristics of the respondent would be burned in my mind forever.	○	○	○	○	○	○	○
	13. I always consider the possibility that if I deny an involuntary civil commitment the respondent may end up committing a violent act in the community.	○	○	○	○	○	○	○
	14. If I denied involuntary civil commitment in a case and shortly afterward, the respondent committed a violent act, I would feel guilty and remorseful about my decision.	○	○	○	○	○	○	○

Appendix N

Regional Differences

This study was the first empirical opportunity to investigate whether judges differed greatly based on the region that he or she worked (see Tables A7 and A8). Results indicated that there were regional differences with regard to the demographic variables of age ($F[3, 388] = 5.743, p = .001$), prior experience as attorney ($F[3, 399] = 40.698, p < .001$), ethnic identity ($\chi^2[12, N = 395] = 29.392, p = .003$), election ($\chi^2[6, N = 396] = 40.515, p < .001$), and education ($\chi^2[12, N = 398] = 122, p < .001$). Judges from the Northeast ($M = 59.79, SD = 7.1$) were significantly older than judges from the South ($M = 54.64, SD = 10.9$). A smaller percentage of Southern judges (56.2%) had previously been attorneys as compared to Northeastern (92.3%), Midwestern (91.1%), and Western (98.2%) judges. A larger percentage of Western judges (28.6%) were appointed and retained via elections as compared to Northeastern (0%), Southern (1.7%), and Midwestern (8%) judges. Northeastern judges (78.5%) were more likely to be elected to the bench as compared to Western (36.6%), Southern (54.5%), and Midwestern (64%) judges. In this sample, Western and Southern judges were more ethnically diverse than Midwestern and Northeastern judges. The South had the highest percentage of African-American judges (10%), while the West had the highest percentage of Hispanic judges (6.3%). All the judges from the Midwest, Northeast, and West had Juris Doctors compared to only 61.2% of Southern judges with Juris Doctors. There were no differences with regard to years on the bench, years with jurisdiction over involuntary civil commitments, and number of civil commitment cases. Chi-square analyses indicated that region was associated with the mode of completing the survey ($X^2 = 18.01, df = 3, p < .001$). Specifically, judges from the South (14%) were less likely to complete this survey online than were judges from the Midwest (33.7%) or Northeast (38.8%). It should be noted that Southern judges were also least likely to be contacted via email; only 23% of Southern judges had available email addresses.

Table A7

Demographic Information by Region – Categorical Variables

Categorical Variables (results in percentage)	Region			
	Northeastern	Southern	Midwestern	Western
Gender				
Male	80%	77.5%	82%	78.6%
Female	20%	22.5%	18%	21.4%
Ethnic Identity				
Caucasian	98.4%	98.4%	96%	88.3%
African American	1.6%	1.6%	1%	2.7%
Latino/Latina	0%	0%	3%	6.3%
Asian	0%	0%	0%	0.9%
Other	0%	0%	0%	1.8%
Education Level				
High School	0%	1.7%	0%	0%
Some College	0%	9.9%	0%	0%
Bachelor Degree	0%	21.5%	0%	0%
Graduate Degree	0%	5.8%	0%	0%
Juris Doctor (JD)	100%	61.2%	100%	100%
Work Status				
Active Judge	89.2%	98.3%	99%	97.3%
Retired Judge	10.8%	0.8%	1%	1.8%
Type of Judge				
Magistrate	0%	31.7%	6.9%	23.3%
Probate	18.5%	14.2%	22.8%	0.9%
Superior	18.5%	1.7%	5.9%	25%
District	18.5%	16.7%	31.7%	38.4%
Circuit	0%	25.8%	25.7%	10.7%
Court of Common Pleas	20%	0%	5%	0%
Other	21.5%	10%	2%	1.8%
Experience with a known false negative decision				
Yes	6.3%	10.3%	7.2%	6.5%
No	93.7%	89.7%	92.8%	93.5%
Election Status				
Elected	21.5%	54.5%	64%	34.8%
Appointed	78.5%	42.1%	28%	36.6%
Appointed and then elected	0%	1.7%	8%	28.6%
Prior experience as attorney				
Yes	92.3%	56.2%	91.1%	98.2%
No	7.7%	43.8%	8.9%	1.8%

Table A8

Demographic Information by Region – Continuous Variables

Continuous Variables [range]	Entire sample	Northeastern		Southern		Midwestern		Western	
	Range	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age	24-76	59.79	7.15	54.64	10.93	57.17	6.7	57.06	7.1
Work Experience									
Years as judge	1-37	13.86	8.29	11.56	7.75	13.32	9.59	12.17	7.44
Years as civil commitment judge	0-35	1.69	8.1	9.77	7.2	10.81	8.95	9.5	6.58
Number of civil commitment cases	0-35000	205.03	589.08	951.97	4345.55	419.97	1605.61	280.98	1468.75