

## GUIDED JURY DISCRETION IN CAPITAL MURDER CASES

### The Role of Declarative and Procedural Knowledge

Richard L. Wiener  
University of Nebraska—Lincoln

Melanie Rogers and Ryan Winter  
City University of New York

Linda Hurt, Amy Hackney, Karen Kadela, Hope Seib, Shannon Rauch,  
Laura Warren, and Ben Morasco  
Saint Louis University

This article analyzes whether state-approved jury instructions adequately guide jury discretion in the penalty phase of first-degree murder trials. It examines Eighth Amendment jurisprudence regarding guided jury discretion, emphasizing the use of “empirical factors” to examine the quality of state-approved instructions. Psychological research and testimony on the topic of the comprehensibility of jury instructions are reviewed. Data from a recently completed simulation with 80 deliberating juries showed that current instructions do not adequately convey the concepts and processes essential to guiding penalty phase judgments. An additional simulation with 20 deliberating juries demonstrated that deliberation alone does not correct for jurors’ errors in comprehension. The article concludes with recommendations for policy and future research.

In October 1993, the Circuit Court for Prince William County, Virginia, tried Lonnie Weeks for the shooting death of Virginia State Trooper Jose Cavazos (*Weeks v. Angelone*, 2000). Mr. Weeks’s uncle was driving a car in which Mr. Weeks was a passenger. Mr. Weeks had stolen the automobile in a home burglary earlier in the month. Trooper Cavazos spotted the car and chased it until it stopped on a highway exit ramp. Weeks stepped out of the vehicle and fired a pistol at the officer, killing him. Mr. Weeks confessed to the homicide and wrote a letter to the jail officer expressing his remorse for his actions (*Weeks*, 2000).

After hearing evidence in the guilt phase, the jury found Weeks guilty of first-degree murder. In the penalty phase, the prosecution attempted to prove two aggravating factors that the jury could use to impose the death penalty, namely, that the defendant was “a continuing serious threat to society” and that the murder was “outrageously or wantonly vile” (*Weeks*, 2000, p. 730). At the conclusion of the 2-day hearing, the jury retired to deliberate on the fate of Mr. Weeks. Apparently the jury was unclear about the meaning of some of Virginia’s penalty phase instructions as evidenced by the fact that they sent the judge two different

---

Richard L. Wiener, Department of Psychology, University of Nebraska—Lincoln; Melanie Rogers and Ryan Winter, Department of Psychology, Graduate Center, City University of New York; Linda Hurt, Amy Hackney, Karen Kadela, Hope Seib, Shannon Rauch, Laura Warren, and Ben Morasco, Department of Psychology, Saint Louis University.

The work reported in this article and the writing of the manuscript were supported by National Science Foundation Grant SBR #9515451.

Correspondence concerning this article should be addressed to Richard L. Wiener, Department of Psychology, 338 Burnett Hall, University of Nebraska, Lincoln, NE 68588-0308. E-mail: rwiener2@unl.edu

questions about the case. The first question was concerned with the possibility of parole following the imposition of a life sentence for first-degree murder under Virginia law. The jury wrote, "Does the sentence of life imprisonment in the State of Virginia have the possibility of parole, and if so, under what conditions must be met to receive parole?" (*Weeks*, 2000, p. 730). The trial judge responded to the jury only by telling them to impose a sentence in accordance with the evidence and not to be concerned with what happens after the state executes the sentence.

The second question, which went directly to the jurors' comprehension of the penalty phase instructions, inquired about the procedure for weighing mitigating and aggravating circumstances under Virginia's approved penalty phase scheme. The question read,

If we believe that Lonnie Weeks, Jr., is guilty of at least 1 of the alternatives, [aggravating factors] then is it our duty as a jury to *issue* the death penalty? Or must we *decide* (even though he is guilty of one of the alternatives) whether or not to issue the death penalty, or one of the life sentences? What is the Rule? Please clarify? (*Weeks*, 2000, p. 730)

The jury appeared to be asking for specific instructions on the procedure that Virginia adopted for weighing aggravating (death penalty enhancing) and mitigating (death penalty diminishing) evidence. The defense attorney asked the trial judge to provide the jury with an explanation of the Virginia balancing rule, requesting that the judge tell the jury that even if it found one or more of the aggravating factors beyond a reasonable doubt, the jury could still impose life in prison as the sentence. Over the defense's objections, the trial judge referred the jury to the specific instructions presented at trial and provided no more guidance. The trial judge commented, "I don't believe I can answer the question any clearer than the instruction, so what I have done is referred them to the second paragraph of instruction number 2 . . .," which reads,

If you find from the evidence that the Commonwealth has proved beyond a reasonable doubt either of the two alternatives, and as to that alternative you are unanimous, then you may fix the punishment of the defendant at death or if you believe from all the evidence that the death penalty is not justified, then you shall fix the punishment of the defendant at life imprisonment or imprisonment for live [*sic*] and a fine of a specific amount, but not more than \$100,000.00. (*Weeks*, 2000, p. 730).

Upon completing its deliberations, the jury did assign the death penalty to Mr. Weeks.

The defense argued on appeal at the U.S. Supreme Court (*Weeks*, 2000, p. 731–732) that this instruction was not adequate because it did not allow the jury to consider the full effect of mitigating evidence in this case. Chief Justice Rehnquist wrote the opinion of the Court in which the justices affirmed the Fourth Circuit's denial of federal *habeas* relief (*Weeks*, 2000). Justice Rehnquist cited the Supreme Court's holding in *Buchanan v. Angelone*, (1998), which held that a state is free to structure how the jury considers mitigation as long as the instructions do not preclude the jury "from considering, and [it] may not refuse to consider, any constitutionally relevant mitigating evidence" (*Week*, 2000, p. 732). In *Buchanan*, the Court upheld the Virginia instruction, the same one that was offered in *Weeks*.

The *Buchanan* Court applied the rule in *Boyde v. California* (1990), and found no “reasonable likelihood that the jury [had] applied the challenged instruction in a way that prevents the consideration of constitutionally relevant evidence” (p. 380). In other words, the case turned on the determination of whether or not there was a reasonable likelihood that the jurors were able to use the instructions presented so that they could consider all of the relevant evidence, hear the mitigating evidence, and reach a just sentencing decision. The remainder of the majority’s opinion presented evidence from the trial to demonstrate that the jury did understand the instruction. Similarly, the dissenting opinion offered evidence that there was a reasonable likelihood that the jury was not able to use the instructions as required by law.

The majority began by noting that there is a presumption that the jury follows the instructions that the court provides (*Richardson v. Marsch*, 1987) and that it understands the judge’s answers to all of its questions (*Armstrong v. Toler*, 1826) (as cited by Justice Rehnquist) (*Weeks*, 2000, p. 733). The majority offered several arguments supporting the presumption that the jury understood the instructions and considered mitigating circumstances. First, the opinion cited the jury poll at the end of the sentence phase, at which time all jurors stated in open court that they agreed with the final sentence. The court reasoned that because all the jurors affirmed that the sentence represented their own conclusions, they indirectly agreed with the stated language that they reached the decision “after considering evidence in mitigation” (*Weeks*, 2000, p. 733). Second, Justice Rehnquist argued that had the jury not understood the trial judge’s instruction, it would have asked further questions, which it did not do. The majority supported this point with the fact that the jury deliberated for two additional hours after asking its second question and before reaching a verdict. The chief justice also made use of the exact language in the jury’s question, “If we believe that Lonnie Weeks, Jr., is guilty of at least 1 of the alternatives. . .” as evidence that it had already found at least one aggravating factor. Justice Rehnquist reasoned from this statement that the jury’s further deliberations were likely devoted to balancing the aggravating and mitigating circumstances. Finally, the majority pointed out that the jury received additional instructions emphasizing its duty to consider mitigating circumstances, first from the defense attorney who told the jury it could use mitigation to reach a life-in-prison sentence even if it found one or more aggravating factors to exist beyond a reasonable doubt. Second, Virginia modified its instructions after it won its appeal in *Buchanan* (1998) to include a definition of mitigating circumstances and a statement telling jurors to consider such evidence before reaching its sentence, “You must consider a mitigating circumstance if you find there is evidence to support it” (*Weeks*, 2000, p. 732).

It is noteworthy that Justice Rehnquist cited the above arguments as “empirical factors” that support the presumption that the jury understood and followed the instructions as offered (*Weeks*, 2000, p. 728). The minority opinion, authored by Justice Stevens, took to task the basis for the empirical conclusions reached by the majority. While the majority offered cogently reasoned arguments, the rebuttal by the four dissenting justices was equally compelling. Justice Stevens claimed that the language in the original instruction could be interpreted to suggest either (a) that life in prison is a valid sentence because of the failure to find an aggravating factor *or* (b) that mitigation justifies such a sentence even if an

aggravating alternative is found. Justice Stevens argued that the question asked by the jury suggests that it did not understand the second procedure (i.e., mitigation offsetting aggravation), which can also lead to a life sentence (*Weeks*, 2000, p. 736). He went on to argue that the reason that the jury failed to question the instruction again was not because it understood it after reading the instruction a second time but rather because, “the jury believed that it would be disrespectful to repeat a simple, unambiguous question that the judge had already refused to answer directly.” (*Weeks*, 2000, p. 736). Indeed it would take a rather exceptionally assertive jury to repeat the question after the trial judge responded to the inquiry. The minority claimed that the verdict forms, none of which offered the jury the option of sentencing the defendant to life in prison after finding an aggravating circumstance beyond a reasonable doubt augmented this confusion. Missing was a form that allowed the jury to sentence the defendant to life on the basis of the offsetting strength of mitigation, despite finding an aggravating circumstance. Stevens argued that the lack of such a form corroborates the misunderstanding that mitigation cannot offset aggravation. Finally, the minority countered that the fact that the jurors affirmed their verdict in open court did not demonstrate that they understood the requirement that they consider mitigation as a procedure that might offset aggravation, it may merely have shown that they all agreed on the sentence because they all found sufficient aggravation to justify the death penalty (*Weeks*, 2000, p. 740).

The importance of the *Weeks* (2000) analysis for social science researchers and commentators is that it demonstrates how the Supreme Court makes decisions using “empirical factors” to examine the presumption that jurors understand and can follow jury instructions. The majority and minority’s analyses rely on the behavior of the jury itself as the most relevant evidence for its understanding or misunderstanding of the instructions. We maintain that comprehension is not simply a property of the jurors who hear the instructions; rather, it results from an interaction between the cognitive and belief systems of the jurors and the structure of the instructions. That is, if researchers provided different types of instructions (e.g., simplified vs. standard instructions) to the same jurors or if they assigned at random different forms of instructions to groups of equivalent jurors (i.e., groups made equivalent by random assignment), comprehension rates would vary as a function of the clarity of the instructions. Theory and research in psychology suggests that there is specific information that instructions are designed to convey and that the information-processing systems of jurors are receptive to that information to a greater or lesser extent, depending on presentation format and content. In other words, an adequate empirical analysis of instruction comprehension must focus not only on the behavior of the jurors but also on the ability of the instructions to convey the substance and procedure of the law. The purpose of this article is to present research results that speak to whether current jury instructions can convey that information and therefore can adequately guide the discretion of jurors in the penalty phase of first-degree murder trials.

In Section I, we trace the development of Eighth Amendment jurisprudence regarding the issue of jury discretion in the sentencing phase of first-degree murder trials. We emphasize the need for consideration of “empirical factors” in the law and the attempt to use empirical reasoning to examine the quality of current instructions. In Section II, we present the findings from psychological

research on the topic of the comprehensibility of jury instructions in first-degree murder cases. We review this research and describe its relevance to the Eighth Amendment concern with guided discretion. More specifically, in this section we trace our collection and application of research data in Missouri and the courts' responses to these and other data offered in postconviction hearings in capital murder cases. In Section III, we present new data from a jury demonstration project. Finally, in Section IV, we review our findings with regard to recommendations for both policy and future research.

### Eighth Amendment Jurisprudence and the Need to Look to "Empirical Factors"

In 1971, the U.S. Supreme Court concluded that the Constitution did not require a state to establish standards or guidelines for juries to use when they were considering a death sentence (*McGautha v. California*, 1971). In other words, the Court held that a jury could act on its own discretion within the broad constraints of law to impose the ultimate criminal sentence. The logic of the Court is best explained by Justice Harlan's observation that to structure such a decision would be so difficult that it might be beyond the ability of human reason (*McGautha*, 1971, p. 204). Justice Harlan made the empirical assumption that jurors would not or could not follow rules designed to guide their discretion in imposing the death penalty and that therefore there was no need to even try and establish a set of guidelines, and of course no need to empirically examine the decision making process of capital jurors.

The influential American Legal Institute (ALI) did not see the death penalty imposition in the same light, as evidenced by the commentary that accompanied the establishment of the Model Penal Code (Anderson, 2000). The ALI in its early meetings concluded that the death penalty was likely to remain part of the U.S. criminal justice system, and therefore it considered seriously the ingredients necessary to regulate and codify its imposition (Anderson, 2000). The ALI focused on the control of jury discretion in invoking capital murder, reasoning that only with controlled discretion could the courts hope to apply the death penalty in a way that was equitable and not capricious. As a result, the ALI in the Model Penal Code (1980, § 210.6) stated, "Those jurisdictions that elect to retain the penalty must confront the special need to provide a fair and rational system of administration and to meet recently developed constitutional standards" (as cited in Anderson, 2000, p. 756). It is with this idea that the criminal justice system first (implicitly) opened the door to considering the empirical basis for juror decisions to impose death. Standards for guiding discretion can be operationalized, and operationalized standards are measurable with methodologies currently available in the psychological sciences. The ALI argued against the assumption that the law could not regulate death penalty judgments, and the Supreme Court was soon to alter its position to agree with the ALI analysis.

In 1972, the U.S. Supreme Court recognized that the death penalty was different from all other punishments and that therefore special consideration for its uniqueness (*Furman v. Georgia*, 1972) should accompany its imposition. Justice Brennan went so far as to say that the death is "a unique punishment in the United States" (p. 286) and that its "finality and enormity" puts it "in a class by

itself.” (p. 289). One legal commentator argued that Brennan’s placement of the death penalty into a class of its own fueled much of the Supreme Court’s judgments about the death penalty in the decades that followed (note, Note, 2001). The Court, having concluded that the death penalty is a fundamentally different type of punishment, had little choice but to establish rules in subsequent cases that acknowledged that difference, treating it with a special set of procedures and processes. In *Furman* (1972), the Court made use of this uniqueness to rule that all existing death penalty schemes were in violation of the Constitution because the imposition of the death was unguided and without a consistent standard. In *Furman*, the Court ruled that the death penalty was cruel and unusual punishment prohibited by the Eighth Amendment to the Constitution. In *Furman* (and other 1976 cases), the Court called for standards to guide juror discretion and with that call made implicit empirical assumptions about inconsistencies in the manner in which jurors decide whether to impose the death penalty. When it appeared that the Supreme Court was interested in the consistency and equity of death penalty imposition, then psychologists devised ways to measure and document jury behavior with regard to its imposition.

In the years following *Furman*, the states revamped their death penalty schemes to guide and structure the decision making process of jurors. In *Gregg v. Georgia* (1976), the Supreme Court upheld Georgia’s death penalty statute, concluding that the death penalty could only be applied if procedures were in place to ensure that the imposition of death was “directed and limited so as to minimize the risk of wholly arbitrary and capricious action” (*Gregg*, 1976, p. 189). In *Gregg*, the Court acknowledged that Georgia’s bifurcation of capital murder trials into a guilt phase to determine whether the defendant was guilty of first-degree murder and a penalty phase, if necessary, to determine whether to impose death as a punishment was an important ingredient to guided discretion demanded by the Eighth Amendment. Further, guided discretion demanded that jury decision making in the penalty phase be aided by aggravating factors listed in statute (those factors that make the defendant worthy of the ultimate punishment) and by mitigating factors also listed in statute (those factors that argued for leniency and a prison sentence instead of the death penalty). The jury should weigh the aggravating and mitigating factors to reach a conclusion about the suitability of the death penalty in a particular case (*Gregg*, 1976, pp. 193–195).<sup>1</sup>

The balancing test offers to psychologists a starting point to operationalize the death sentence process. It suggests that researchers might measure the relative endorsement of aggravating factors, mitigating factors, and the weights that jurors assign to these factors to capture and document the jurors’ sentencing process. Although the task is conceptually simple, it proves to be much more difficult in practice because of the richness of the evidence and the influence of deliberations in jury trials. Despite the difficulty of the task, the Supreme Court inadvertently opened the door to empirical investigation of the penalty phase when it endorsed

---

<sup>1</sup>In *Zant v. Stephens* (1983) the Supreme Court also held that state appellate review in death penalty cases was automatic so that the guided discretion that jurors were required to practice would be evaluated separately by the state’s appellate court system in each case. We do not review this line of cases, despite their obvious importance to capital murder, because they do not bear directly on the use of empirical evidence to evaluate jury discretion.

the balancing system offered in Georgia's death penalty statute. It is undeniable that the Court did not do so with the intention of relying on social science data in its capital murder decisions. In fact, the Court rejected these data when it considered challenges to the death qualification process (*Lockhart v. McCree*, 1986) and claims of racial discrimination in the imposition of the death penalty in Georgia (*McClesky v. Kemp*, 1987). Nonetheless, the holding in *Gregg* (1976) was instrumental in launching psychological investigations of the penalty phase, which we take up in the next section in detail.

In subsequent cases, the Supreme Court refined the use of mitigation in the penalty phase of capital murder trials. In *Lockett v. Ohio* (1978), it offered a broad interpretation of mitigation content when it held that jurors may consider "any aspect of a defendant's character or record and any of the circumstances of the offense that the defendant proffers as a basis for a sentence less than death." (p. 604). Further, in *Mills v. Maryland* (1988) the Court invalidated Maryland's jury verdict forms because they did not allow a prison sentence when all jurors found different mitigating factors that were sufficient to offset aggravating factors. It is the burden of the prosecution to prove at least one aggravating factor to all jurors. As in all other criminal cases, the prosecution must prove its case, including the existence of aggravating circumstances sufficient to warrant the death penalty, beyond a reasonable doubt (*Cage v. Louisiana*, 1990; *In re Winship*, 1970). However, mitigation need not be demonstrated beyond a reasonable doubt, it merely needs to be demonstrated to the satisfaction (presumably by the preponderance of the evidence) of the individual juror. Further, following *Mills* (1988), all jurors do not need to find the same mitigating circumstance to outweigh the aggravating evidence, each juror need only find some mitigating factor(s) that offsets the aggravating factors in order for the jury to find the defendant not worthy of the death penalty.

In its most recent death penalty cases, the Supreme Court further clarified the penalty phase elaborating on its two distinct processes. In *Buchanan* (1998) the Court spoke of the "eligibility phase and the selection phase" (p. 761). In the eligibility phase, the jury determines whether a defendant found guilty is eligible for the death penalty, most often by considering the existence of statutory aggravating circumstances. In the selection phase, the jury determines whether to impose the death penalty upon a particular defendant (i.e., select the eligible defendant for either a prison sentence or the death penalty). States must structure the selection phase so that "the sentencer may not be precluded from considering, and may not refuse to consider, any constitutionally relevant mitigating evidence" (p. 276). However, the state is not required to "structure in a particular way the manner in which juries consider mitigating evidence." (p. 761). The state need not list specific mitigating factors, nor must it provide a conceptual description of mitigating circumstances. Thus, *Buchanan* reinforced the existing law, extending the wide latitude of jury discretion in the selection phase as long as the jury is not precluded from considering any permissible evidence in mitigation. Indeed, the Supreme Court found the Virginia instructions passed Constitutional muster even though those instructions only indirectly told jurors to consider mitigation after they unanimously found one of the statutory aggravating factors to exist beyond a reasonable doubt. The exact language was "if you believe from all the evidence that the death penalty is not justified, then you shall fix the punishment of the

Defendant at life imprisonment” (*Buchanan*, 1998, p. 273). The Court held that the Virginia instructions do not preclude the jury from considering all mitigating circumstances allowable by the Constitution.

Even after *Buchanan* (1998), the *Boyd* (1990), rule remains; that is, there cannot be a “reasonable likelihood that the jury has applied the challenged instruction in a way that prevents the consideration of constitutionally relevant evidence” (p. 380). The fact that the Supreme Court looked to “empirical factors” in the subsequent case (*Weeks*, 2000; see above) suggests that empirical analysis is relevant to this issue. Given that the Court itself has examined empirical evidence to determine whether jurors have understood the instructions in the penalty phase, it should come as no surprise that psychologists have also attempted to challenge jury instructions with empirical evidence that demonstrates high levels of confusion under state-approved death penalty phase instructions. Consider first an Illinois challenge concerning the comprehensibility of its instructions.

In the first-degree murder case of *Free v. Peters*, (1991) the late Professor Hans Zeisel conducted a jury comprehension survey to determine whether jury-eligible citizens were able to understand the directions offered in the 1987 Illinois Pattern Instructions (IPI) for capital cases. Zeisel presented a fact pattern describing the homicide in *Free v. Peters* (1991) and a copy of the instructions used at Free’s trial to people reporting to jury duty in Illinois. Zeisel asked the respondents a series of multiple-choice questions to ascertain their understanding of the IPI instructions. Zeisel found high levels of miscomprehension, especially with regard to the respondents’ understanding of the appropriate use of mitigation. Up to 46% of the respondents answered wrong some basic questions regarding the use of mitigation in the penalty phase of the trial (*Gacy v. Wellborn*, 1993).

Magistrate Judge Weisberg reviewed Zeisel’s data and concluded that

there is a reasonable likelihood that a substantial number of jurors who received the IPI instructions and [a substantial number] of Free’s jurors believed that only the statutory mitigating factors, or factors comparable to them, could preclude the imposition of the death penalty

and that

The Free jury, like juries receiving the IPI instructions, was probably confused about which side, if any, had a burden of persuasion and what the nature of that burden was. (U.S. ex rel. *Free v. Peters*, 1991, p. 708)

Based on *Boyd* (1990), the Court concluded there was a “reasonable likelihood that the jury has applied the challenged instruction in a way that prevents the consideration of constitutionally relevant evidence” (p. 380). The jury was not able to use mitigation in a way that that allowed them full consideration of all potential mitigating factors.

However, in *Gacy* (1993), the Seventh Circuit Court of Appeals overturned the lower court findings and denied *habeas corpus* relief to Free and to serial killer John Wayne Gacy. The Seventh Circuit held that although Zeisel’s data may show that the IPI do not produce ideal juror comprehension, they fail to show that an improvement is possible with alternative forms of jury instructions. The absence

of a control group supplied with simplified instructions prevents the conclusion that there are alternative ways to instruct the jurors that prevent the high levels of confusion that Zeisel measured. The failure to include such a control group was a fatal flaw in Zeisel's study according to the Seventh Circuit (*Gacy*, 1993, p. 311). In *Free v. Peters* (1991) Judge Posner of the Seventh Circuit Court of appeals writing for the majority added that Zeisel's work did not document a reasonable likelihood that the *Free* jurors misunderstood the IPI penalty phase instructions because of the failure of that work to adequately simulate a jury evaluating instructions. Quoting Judge Posner,

The tests have, however, two fatal flaws. The first is lack of comparability between the test setting and the setting of the sentencing hearing. The second is the lack of a control group consisting of persons administered a test containing what Zeisel (or *Free*'s lawyers) would consider adequately clear instructions. . . (p. 705)

Even if the problem of comparability were not fatal, the absence of a control group would be. And it was such an easy problem to solve. All Zeisel would have had to do was administer his test to another randomly selected group of "jurors," only this time altering the instructions so that they were clearer. (p. 706)

### Psychological Research and Guided Jury Discretion

Wiener, Pritchard, and Weston (1995) set out to address the control group problem that the Seventh Circuit adroitly identified. Using the Missouri Approved Instructions (MAI) for the penalty phase in capital murder trials, they measured the comprehensibility of sentencing instructions and explored the utility of psycholinguistic interventions to improve juror understanding. In two experiments, they presented to 175 Missouri citizens willing to impose the death penalty four different sets of penalty phase jury instructions: (a) *control condition*: The instructions included only an outline of the sentencing task, leaving out definitions of important elements (e.g., aggravation, mitigation, and beyond a reasonable doubt); (b) *original MAI*: The original MAI instructions consisted of the actual instructions that were used in the trials of *Brown v. State* (1989) and *Butler v. State* (1990); (c) *revised MAI instructions*: The revised instructions consisted of the MAI penalty phase instructions adjusted to allow the sentencer to consider mitigating circumstances even when they are not agreed upon by all jurors (*Mills*, 1988); and (d) *model instructions*: The model instructions were rewritten in language simplified and clarified in accordance with the findings of recent psycholinguistic research in jury decision-making.<sup>2</sup>

Using two independent fact patterns from recent Missouri trials, Wiener et al. (1995) presented to jury-eligible, "death-qualified" participants (i.e., those willing to impose the death penalty) fact summaries, written versions of the instructions (which were also read aloud on a tape by an ostensible judge), and a jury comprehension questionnaire. The comprehension inventory that the participants completed after reviewing one set of facts and the accompanying instructions

---

<sup>2</sup>See Table 4 for examples pattern, baseline, and model (simplified) instructions from the most current simulated jury study described in this report. Modifications to these instructions were based on recent psycholinguistic research findings (Charrow & Charrow, 1979; Elwork & Sales, 1985; Elwork, Sales, & Alfani, 1982; Severance, Greene, & Loftus, 1984; Strawn & Buchanan, 1976).

consisted of 29 questions that measured the jurors' understanding of how to apply the law in accordance with each of five areas of recent Eighth Amendment jurisprudence.

Analyses of both data sets showed that the respondents were quite confused about the meaning of the instructions. The upper boundaries of confidence intervals for knowledge of all scales except two failed to reach the 70% accuracy level. Comparing the combined MAI conditions (original and revised) to the control condition showed that comprehension of the reasonable doubt standard of evidence, the Missouri sentencing procedure, and juror agreement about mitigation was not significantly greater than baseline knowledge in the *Brown* (1989) fact pattern, a heinous and emotional homicide involving an adult sexually abusing and strangling a little girl. Moreover, in the *Butler* fact pattern, a less heinous crime involving a husband shooting his wife for insurance money, these same analyses showed that MAI instructions were not significantly better than baseline for any of the comprehension scales. Further, with the *Butler* data, but not the *Brown* data, the model instructions produced greater comprehension relative to the MAI conditions with regard to the Missouri sentence procedure, juror agreement about mitigation, and total comprehension averaging all the scales.

Thus, in response to Judge Posner's attack on the Zeisel study, Wiener et al. (1995) were able to show that pattern instructions, simplified according to principles of psycholinguistics, can improve jury comprehension relative to a baseline control. It is not simply the case that comprehension falls short of an ideal (see *Gacy*, 1993); rather, it falls short of comprehension levels attainable with simpler semantic and grammatical composition. Perhaps most important, when the researchers divided the respondents (across cases) into those who sentenced the defendant to death and those who did not, they found that across both fact patterns, those who assigned the death penalty showed poorer understanding of the reasonable doubt standard, Missouri sentence procedure, appropriate mitigation content, and total comprehension. These correlational data suggest that jurors who fail to understand the instructions are more likely to assign the death penalty, at least without the aid of jury deliberations.

In a follow-up study, Wiener et al. (1998) applied social cognition theory to try to understand why jurors performed so poorly on comprehension tests. They presented the same *Brown* (1989) and *Butler* (1990) fact patterns to 212 death-qualified undergraduate mock jurors (i.e., those willing to impose the death penalty but not those who automatically impose it). Wiener et al. presented the *Brown* case (i.e., written fact pattern, instructions, and survey) to half of the jurors and the *Butler* case to the other half. In this study, all participants read a copy of the actual instructions used at trial while listening to a tape of an ostensible judge reading those MAI instructions aloud. These were the same instructions used in the actual MAI instruction conditions in Wiener et al. (1995).

Respondents answered a jury comprehension survey as they did in Wiener et al. (1995); however, the construction of the survey reflected some more subtle discriminations based on legal and psychological distinctions. In order to process case facts in accordance with the law, jurors must construct a theory of sentencing, which is mostly based on the jurors' interpretations of the language that the judge uses when the judge directs jurors during the guilt and penalty phases (Pennington & Hastie, 1986, 1988). However, following Pennington and Hastie

(1986), we hypothesized that the process of constructing the law from the judge's instructions is an active one that also relies heavily on the juror's own cognitive system. We reasoned that mock jurors fail to understand jury instructions, in part because of the content, processes, and resulting biases in their own information-processing systems.

Because death penalty jurors make sentencing decisions usually only one time in their lives, they do not gain independent knowledge of the process from practice. As a result, they have limited opportunity to develop the cognitive skills that they need to apply the law correctly. E. R. Smith and colleagues (E. R. Smith, 1988, 1990; E. R. Smith & Branscombe, 1987, 1988; E. R. Smith, Branscombe, & Bormann, 1988; E. R. Smith & Lerner, 1986; E. R. Smith & Zarate, 1990) distinguish between *declarative knowledge* (i.e., meaning and content based information stored as semantic concepts, schemata, scripts, or prototypes) and *procedural knowledge* (i.e., the active component of the cognitive system that operates on the declarative knowledge stored in long and short term memory) of people and social events. E. R. Smith et al. (1988) demonstrated that people who make repeated social judgments learn to apply generic rules, which are related to specific content even when the rules, to some extent, function independently of that content. However, E. R. Smith (1989) found that practice effects are not entirely content independent. Procedural practice not only improves the individual's ability to enact cognitive procedures but also makes more accessible in memory abstractions (e.g., schemata) related to the content of the practice.

The implications of this work for juror decision making in capital murder trials are significant. To apply the law and balance aggravating and mitigating circumstances, it is not sufficient for jurors to represent the meaning of the law in their constructed theories of sentencing. Without adequate declarative and procedural knowledge of the decision process, jurors will be unable to accurately understand and apply the law. We constructed this and subsequent jury comprehension surveys with a flowchart model of the approved instructions for the penalty phase of a capital murder trial. We diagrammed the instructions, parsing the language into concepts and steps that are necessary to reach a sentence in the penalty phase of the trial. (See Figure 1.)

We treated each term of law (e.g., *aggravating circumstances*, *beyond a reasonable doubt*, and *mitigating circumstances*) as a declarative concept and each of the decision processes implied in the flowchart as a procedural rule (e.g., all jurors must find at least one [the same] aggravating circumstance before considering the death penalty and going on to examine mitigating circumstances) (Wiener et al., 1998). Some of the concepts and rules appear directly in the state-approved instructions and pertain specifically to state statutes (e.g., *beyond a reasonable doubt* means to be firmly convinced but not convinced beyond all doubt of any kind). Other concepts and rules arise in constitutional law (e.g., a juror may consider any aspect of the crime or the defendant's life as a mitigating circumstance). From the resulting typology, we created independent scales that measured four types of knowledge (i.e., declarative state law, procedural state law, declarative constitutional law, and procedural constitutional law) and wrote multiple-choice questions to measure each scale separately. Multiple items (four scales with 65 questions in total) measured each concept (Wiener et al., 1998).

Research participants completed the resulting jury comprehension survey

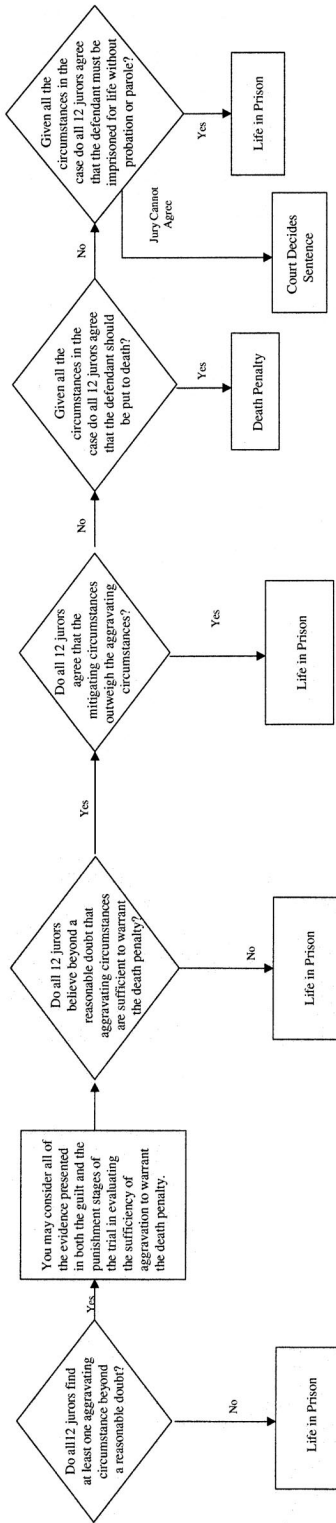


Figure 1. Flowchart of Missouri Approved Jury Instructions, in the penalty phase of capital murder trials, from *State v. Kneese* (1999).

while in possession of the printed penalty phase MAI instructions used in the actual trials. An example of a declarative state item was

After reviewing all the evidence in the case, a juror found only one mitigating circumstance to be true. The remaining jurors do not agree that this specific mitigating factor was true. Does the law prohibit this juror from considering this mitigating factor? (Wiener et al., 1998, p. 131)

The available answers were *yes*, *no*, and *do not know*. Similar to this question, each of the items had a correct answer (for the above question, *no*). For each scale, we calculated an accuracy score that controlled for guessing by assigning a value of 1 to the correct answer,  $-1$  to an incorrect answer, and 0 to the *do not know* response. We converted the results to percentages by summing the scores in each scale, dividing by the number of items in the scale, and multiplying the result by 100. In this manner, a perfect score for each of the scales was 100%, a completely incorrect score was  $-100\%$ , and a chance score (reached by guessing or answering *do not know*) was 0% (Wiener et al., 1998).

In the *Brown* (1989) and *Butler* cases, all scales exceeded zero but the overall mean scores were low, ranging from .067 for the state procedural knowledge to .299 for the state declarative knowledge (Wiener et al., 1998). Knowledge of all types of information was poor. Across both cases, the procedural knowledge was less well comprehended than was declarative knowledge, especially the state law procedural knowledge. That is, respondents were most confused with rules that guided them in applying the law to the facts of the cases (e.g., how to balance aggravating and mitigating circumstances and not to proceed to mitigating factors unless an aggravating circumstance is agreed upon).

Wiener et al. (1998) created a measure of sentence certainty by multiplying respondents' sentences (1 = *life in prison*;  $-1$  = *death penalty*) by their certainty (1 to 9) in the sentence. Sentence certainty was a continuous variable ranging from 1 (*very certain of the death penalty*) to 9 (*very certain of life in prison*). We found that both strength of aggravation, dichotomized into "strong" and "weak" for the sample, and strength of mitigation, also dichotomized into strong and weak in the sample, exerted main effects on sentence certainty (i.e., weak aggravation and strong mitigation were associated with greater certainty in a life sentence). However, if subjects followed the procedures in the MAI, the effect of mitigation strength should only have been emerged under conditions of strong aggravation. Jurors should only consider mitigation when aggravation is strong; a life sentence should result with weak aggravation. We found no evidence for the interaction of strength of aggravation and strength of mitigation as the instructions demand.

Finally, when we regressed sentence certainty on the four scales of comprehension—state declarative, state procedural, constitutional declarative, and constitutional procedural law—we found an effect only for the constitutional procedural questions. Those mock jurors who understood the least about constitutional procedural issues were most certain in sentencing the defendant to death. Those who showed the greatest knowledge on this scale favored a life-in-prison sentence. In summary, the Wiener et al. (1998) study demonstrated that jury-eligible citizens do not comprehend approved jury instructions very well (especially procedural information) and, to the extent that they do not understand procedures, that they fail to follow the law and tend to favor the death penalty.

The Missouri courts have considered the data from these studies in the postconviction hearings of a number of capital murder cases in which Richard L. Wiener testified specifically about the findings of Wiener et al. (1995) and about some of the results in Wiener et al. (1998). The data came to the attention of the Missouri Supreme Court in *State v. Jones* (1998) as part of a claim that the counsel in the Jones case was ineffective because the attorneys failed to offer evidence that jurors do not understand the MAIs despite the fact that these data were available at the time of the Jones trial. The Missouri Supreme Court dismissed the claim of ineffective assistance of counsel, simply stating, "This research does not necessarily support a conclusion that the jurors in this case were unable to understand the MAI instructions" (p. 181). The Missouri Court gave no other reason for rejecting these data.

However, in 1999 the Missouri Supreme Court again considered these studies in *State v. Deck* (1999). Mr. Deck, convicted of two counts of first-degree murder, claimed that he was denied his right to be free from cruel and unusual punishment as required by the Eighth Amendment to the U.S. Constitution and Article I, Sections 10 and 21, of the Missouri Constitution. The Missouri Supreme Court dismissed Deck's motion, finding fault with the adequacy of the jury simulation used in the Wiener et al. (1995, 1998) studies. In psychological terms, the Court complained that the research methodology used in these studies was weak with regard to external and construct validity. Because the simulation did not use deliberating juries presented with realistic trial materials, the studies failed to measure juror comprehension of the instructions in a relevant manner. Specifically, the opinion stated,

Dr. Weiner's study, however, must be discounted because the people interviewed for the study did not act as jurors. They were given hypothetical facts that were different than the facts in this case, and they did not hear the testimony of witnesses, observe physical evidence or deliberate with eleven other jurors. (*State v. Deck*, 1999, p. 542)

Similar to the *Free v. Peters* Court, the *Deck* (1999) Court found that the simulation procedure was inadequate to support the conclusion that pattern instructions confuse jurors and prevent them from exercising the guided discretion required by the prohibition to cruel and unusual punishment. Going further, the Court found that no reasonable person could be confused about mitigation because Missouri instructions always present it in contrast to aggravation; therefore, taking the instructions as a whole, jurors always see mitigation as the opposite of aggravation.

In Section III, we present some empirical data to address the external and construct validity issues that the Seventh Circuit and the Missouri Supreme Court raised. First, we review the empirical research that considers the claim that taken in their context, jury instructions are comprehensible to a reasonable person. Several psychological investigations now exist that challenge the notion that jurors comprehend penalty phase instructions. In early work, psychologists demonstrated that individuals qualified to sit as jurors poorly comprehend jury instructions pertaining to a variety of procedural and substantive content areas (Charrow & Charrow, 1979; Elwork & Sales, 1985; Elwork et al., 1982; Severance et al., 1984; Strawn & Buchanan, 1976). However, work that is more recent

focuses directly on the issue of juror comprehension of instructions given in the penalty phase of capital murder trials. For example, Luginbuhl (1992) used a mock jury paradigm that allowed participants to deliberate after considering the facts and law of the case and found that eligible jurors' understanding of the definition and use of mitigating circumstances as explained in North Carolina's death penalty instructions was severely limited. However, instructions modified to be consistent with North Carolina Supreme Court holdings, which require explanations of mitigating circumstances and final sentencing, significantly improved mock jurors' comprehension. In a follow-up study with a community sample, Blankenship and Luginbuhl (1994) found that mock jurors who showed the greatest comprehension of penalty phase instructions, especially with regard to mitigating instructions, would have been ineligible to serve on a jury because they were not "death qualified" (i.e., they were unwilling to impose the death penalty in any case). Further, the jurors with the highest level of comprehension were the least accepting of the prosecution's arguments about aggravating circumstances. In short, Luginbuhl and colleagues demonstrated that juror miscomprehension of mitigation in penalty phase instructions is consistent, it may bias a juror against the defense, and careful rewriting of instructions can reduce miscomprehension.

Work by Haney and colleagues (Haney & Lynch, 1994; Haney, Sontag, & Costanzo, 1994) also demonstrated that jury instructions are difficult to understand. Haney and Lynch (1994) asked college student mock jurors to define aggravating and mitigating instructions after reviewing California's penalty phase instructions (which did not contain explicit definitions of either aggravation or mitigation) and found that only 64% were even partially correct in their definitions. Perhaps more important, fewer (only 47%) could provide a partial definition of mitigating circumstances. Finally, to make matters even worse, many of Haney et al.'s participants confused specific aggravating and mitigating factors; for example, 25% inaccurately classified acting under the influence of mental or emotional disturbance as an aggravating factor. This is a statutory mitigating factor under California law. Haney et al. (1994) extended these findings in an interview study of former death penalty jurors in California and found that they too transformed information introduced as mitigating into aggravation. For example, a number of former jurors thought that a defendant's pro-social childhood conduct, introduced as mitigation, was actually aggravating because it was a reason to hold the defendant to a higher set of moral obligations than is required of others without such a background. In other words, the jurors did not simply discount pro-social conduct as a mitigating factor, instead they treated it as a statutory aggravating factor, even though it was not such a factor as defined in California law.

After California modified its instruction to include definitions of mitigation and aggravation, Haney and Lynch (1997) tested the new instructions with a sample of undergraduate mock jurors and found significant improvement for understanding aggravation. With the modified instructions, 71% of respondents were at least partially correct in their understanding of aggravation. However, the new instructions did little to improve the jurors' understanding of mitigating circumstances; only 51% were able to provide even a partially correct definition. Most recently, Lynch and Haney (2000) presented a videotaped reenactment of a murder trial to a community sample of mock jurors and once again found poor

overall comprehension of instructions. Most significantly, Lynch and Haney (2000) presented videotapes that varied the race of the defendant so that some participants viewed an African American defendant and others viewed a European American defendant. This research presents the disturbing finding that the African American defendant was more likely to receive the death penalty than was the European American defendant, despite the fact that the rest of the facts in the case were identical. Particularly troubling for the purposes of the current article was the fact that the differential racial sentencing among mock jurors was the greatest among those who scored low, as opposed to high, on the comprehensions of penalty phase instructions. Apparently, when left to their own inadequately guided discretion, mock jurors are likely to rely on their own racial prejudices when invoking punishment for first-degree murder.

In summary, we conclude that the research in the Luginbuhl, Haney, and Wiener research laboratories suggests that the courts should be cautious in concluding that reasonable people understand mitigation and aggravation as presented in pattern instructions. The data collected to date points out a number of ways in which jurors can (and do) misunderstand the basic concepts that make up penalty phase instructions. Taken as a whole, the existing literature converges on a serious challenge to the assumption that reasonable individuals understand jury instructions. Although the studies reviewed so far raise a “reasonable likelihood that [juries have applied] the challenged instruction in a way that prevents the consideration of constitutionally relevant evidence” (Boyde, 1990, p. 380), our current research in this area goes even farther to challenge the assumption that juries follow the instructions as provided (Richardson, 1987).

### A Psychological Investigation of “Empirical Factors”

Beginning in 1998 and continuing until the present, the psychology and law laboratory (first at Saint Louis University and currently at the University of Nebraska at Lincoln) conducted a series of investigations to learn how jurors reach penalty phase judgments in first-degree murder trials. Our research team took as its point of departure Missouri homicide law and the instructions presented to jurors sitting on first-degree murder trials in that state. The questions that this work asked evolved from the reaction of the courts to psychological data and the potential answers to these questions offered by theory in psychology.

#### *Interview Study*

Our work adopted Pennington and Hastie’s (1986, 1988) three-stage, story construction model of jury decision making, which other jury researchers (V. L. Smith, 1991, 1993; Wiener, Wiener, & Grisso, 1989) and social science and law commentators (Hans, 1988; Wiener, 1993) have strongly endorsed. In the first stage, jurors collect the information presented by trial witnesses and attorneys and modify it according to their own general knowledge of the social and physical world to construct their own stories of the case. Next, jurors form composite theories of the law that integrate the judge’s instructions with their own lay models of moral and legal justice. As a result of this process, the final trial stories and legal theories are likely to include accurate representations, unintended misrepresentations, and perhaps purposeful distortions (V. L. Smith, 1991). In the

final stage, jurors apply their new theories to interpret their reconstructed stories and reach their final decisions.

Although researchers have applied the story construction model only to study juror verdicts (Pennington & Hastie, 1986; V. L. Smith, 1991, 1993; Wiener et al., 1989), it is also relevant to capital murder sentencing decisions. We reasoned that death penalty jurors construct homicide stories and defendant life histories, combining trial evidence and their own representations of social reality. Listening to the judge's penalty phase instructions, the jurors construct their own theories of law by modifying the judge's directions to fit their own theories of homicide, murderers, victims, and legal culpability. Finally, jurors apply their constructed theories to the trial stories to reach sentence decisions. Deliberation will modify the conclusions only if it alters the jurors' constructed stories or theories of law.

Work by V. L. Smith (1991, 1993) and by Wiener et al. (1991) supports the notion that people rely heavily on their own general knowledge of social reality, that is, their stored declarative knowledge about the law when they process trial-like information. V. L. Smith (1991) measured readily accessible cognitive prototypes about crime and showed that people use these prototypes to make determinations of crime category membership. Most important, V. L. Smith (1991) found that neither presenting preliminary instructions (i.e., regarding the burden of proof, the standard of proof, and the role of evidence in reaching verdicts) nor preliminary instructions plus crime definitions from the Illinois Pattern Jury Instructions (IPJI) improved the accuracy of participants' convictions for a number of crime scenarios. Although each scenario matched one of the IPJI crimes, participants were still most likely to convict when the facts matched most closely their own schemata of assault, burglary, or kidnapping. In a second set of studies, V. L. Smith (1993) instructed mock jurors not to rely on inaccurate features of their shared kidnapping schemata. Warning jurors against the use of specific inaccurate features reduced the likelihood that they would convict when the fact pattern matched those that they thought to be typical of kidnappings.

In general, we hypothesized that mock jurors fail to understand the law, in part because of the content, process, and bias in their own information-processing systems. It follows that rewriting jury instructions in clear and simple language will not completely de-bias juror decisions. We reasoned that only by instructing jurors specifically not to rely on faulty declarative knowledge about the law could we improve comprehension of the actual instructions delivered at trial. Therefore, we set out to catalogue the types of errors that people share about the sentencing process articulated in the law of first-degree murder.<sup>3</sup> We began this process by conducting a set of open-ended interviews with 126 jury-eligible citizens living in and around the St. Louis area.

---

<sup>3</sup>Another purpose of the open-ended interviews was to document and classify the types of murder stories that jury-eligible citizens hold about first-degree murder. We expected that the types of stories that people know and tell would, in part, determine the types of sentences that they assign to new cases. Story data from this investigation are reported elsewhere. The interested reader is referred to Wiener, Richmond, Seib, and Rauch (2002), which describes the types of murder stories that we found to be prevalent in the interview sample.

### *Interview Method*

*Research participants.* Participants were 63 male and 63 female death-qualified, jury-eligible citizens in the St. Louis area. Death-qualified jurors are willing to invoke the death penalty but do not automatically do so. The mean age of the participants was 40.6 years. The sample, selected to be representative of the greater St. Louis metropolitan area in demographics, consisted of 70% Caucasians, 22% African Americans, 3% Hispanic, and 5% other. The modal respondent had finished high school and attended some college (but did not graduate), and 80% were employed full or part time.

We recruited these individuals through local newspaper advertisements inviting participation in a study investigating “how people think about homicide.” Interested respondents called a local answering service and left their names and telephone numbers. A research assistant telephoned candidates and screened them for jury eligibility. We identified respondents who were 21 years of age, possessed a valid Missouri driver’s license, and/or were registered voters in the state as potential participants for the interview. The state of Missouri uses these criteria to select people for jury service. In addition, we administered our “Attitude Toward the Death Penalty Questionnaire” over the telephone. It consists of four statements regarding attitudes about the death penalty which allowed the participants to indicate that they (1) “would always vote to sentence a guilty defendant to death”; (2) “would not necessarily vote for the death penalty in every case where the law allowed it . . . but would consider the facts of the particular case. . .”; (3) “have doubts about the death penalty, [but] would be able to . . . vote for a death sentence where the law allowed it. . .”; or (4) “have such strong doubts about the death penalty that they would be unable to find the defendant guilty and vote for a death sentence where the law allowed it. . .” Participants who answered *yes* to Statement 1 or 4 were considered not to be death qualified and were judged ineligible to serve in capital murder trials. We did not ask those respondents to participate further in this research; instead we invited eligible participants to come to campus and participate in the interviews. When they arrived at the laboratory, participants completed a written questionnaire once again asking about their sentiments toward the death penalty. In accordance with the current standard (*Wainwright v. Witt*, 1985), we also asked each respondent to indicate

Which of the following statements best describes how you would conduct yourself as a juror on a capital murder case? (a) I have such strong sentiments about the death penalty that *they would* seriously affect me as a juror and would prevent or substantially impair my performance in accordance with my instructions and oath, or (b) My sentiments about the death penalty *are not so strong that they would* seriously affect me as a juror and would prevent or substantially impair my performance in accordance with my instructions and oath.

All 126 respondents circled the second option providing further support that they would be eligible to serve on a capital murder case.

*Materials and procedure.* One of four trained research assistants interviewed each of the first set of 76 participants (recording answers to open-ended questions on paper and on audiotape). The interviewers asked the participants to describe in detail a first-degree murder story in which a jury justly convicted a

perpetrator of committing first-degree murder. The stories were composed of material that could have been recalled from personal experiences or secondary sources such as media accounts. The data from these stories, which measure the existing stories that people use to think about murder, appear elsewhere (Wiener et al., 2002), and we do not consider them further in this report. However, part of the purpose for asking the participants to describe a first-degree murder case was to assist the participants in making available their recollections and understandings of the law of homicide. Next, we asked the participants to list up to 10 attributes of the defendants and victims in the murder stories that they had brought to mind. Following their list of defendant and victim attributes,<sup>4</sup> participants were read the following instructions:

In evaluating the appropriate punishment in first-degree murder cases, the courts ask jurors to consider both aggravating and mitigating circumstances. Jurors decide upon punishments after they find the defendant guilty of first-degree murder. 1) Do you know what “*aggravating circumstance*” means as it is used in the law? If you do, please tell me, in your own words what the phrase means. If you do not know what it means, please tell me what you think the phrase means.

After recording the response, the interviewer provided a definition of *aggravating circumstance* consistent with Missouri law: “Any factor about the crime or the defendant that makes the defendant more deserving of the death penalty.”

Next the interviewer asked,

2) Do you know what “*mitigating circumstance*” means as it is used in the law? If you do, please tell me, in your own words what the phrase means. If you do not know what it means, please tell me what you think the phrase means.

Once again, the interviewer recorded the answer and then presented a definition consistent with Missouri law: “Any factor about the crime or the defendant that makes the defendant less deserving of the death penalty.” We asked each question in the order in which the issues appear in the Missouri Pattern Jury Instructions for the penalty phase of capital murder. We provided brief definitions for the concepts because later items (i.e., listing 10 aggravating and 10 mitigating factors) required that participants possessed a working understanding of these elements of law.

However, recognizing that providing the answer to the aggravating question might trigger an answer to the mitigating answer, we collected written survey data on 50 additional respondents (i.e., rounding out the 126 interviews). These respondents began by writing out scenarios describing first-degree murder stories that came to mind when we asked about them. Next, respondents wrote out definitions of *aggravating circumstance* and *mitigating circumstance*, but this time, in counterbalanced order, half defined *aggravation* first and half described *mitigation* first. We used the second sample to control for any possible order

---

<sup>4</sup>These data were used to create concept maps of typical victims and defendants in capital murder. The results do not pertain to the data reported here and are therefore not considered further in this article.

effects in the types of errors that naive respondents might make about the meaning of aggravating and mitigating factors.

### Results

*Coding responses.* We developed coding categories for aggravating and mitigating answers that the interviewees provided both in oral and written format. Research assistants established the categories after they read through the participants' responses to both the aggravating and mitigating questions. Table 1 provides a list of categories for aggravating factors, and Table 2 provides a similar list for mitigating factors.

Four independent raters coded a sample of 15 protocols to establish reliability levels for the coding scheme. For aggravating factors, percentage agreement ranged from 81% to 85% between coders, with the mean rate of agreement equal to 83%. For mitigation, agreement across coders ranged from 79% to 83%, with the mean equal to 82%. We concluded that our system of coding reliably captured the types of responses that the participants made when they defined aggravating and mitigating circumstances. Further, comparisons between the original 76 respondents and the second 50 showed no significant differences in the number of responses coded in categories in aggravation or mitigation (see Table 3). Therefore, our discussion of the results of the interview study is collapsed across both samples because the way in which the responses were collected failed to influence the distribution of the responses across categories.

*Definitions.* In Table 1, we grouped the aggravating factor definitions into three categories. Absent definitions or correct definitions present little difficulty for instructing jurors (e.g., "Circumstances or events which would cause a juror to decide in favor of the death penalty in a first-degree murder"). Jurors who begin

Table 1  
*Aggravating Circumstance Definitions Among Uninstructed Jury-Eligible Citizens*

Factor	%
Aggravating factors that are without definition or correct	
Provides no definition or an incoherent definition	18
Makes the defendant more deserving of harsh punishment	4
Aggravating factors that are wrong or too narrow	
Proves the defendant more likely to be guilty of first-degree murder	10
Suggests that the defendant was malicious but fail to describe the implications for responsibility or punishment	3
Suggests that the crime was heinous but fail to describe the implications for responsibility or punishment	12
Aggravating factors that imply less punishment	
Makes the defendant less deserving of harsh punishment	2
Proves the defendant less likely to be guilty of first-degree murder	3
Caused the murderer to kill the victim but that do not constitute direct provocation or aggravation	18
"Aggravated" the defendant and/or provoked the defendant to kill the victim	29

Table 2  
*Mitigating Circumstance Definitions Among Uninstructed Jury-Eligible Citizens*

Factor	%
Mitigating factors that are without definition or correct	
Provides no definition or an incoherent definition	31
Makes the defendant less deserving of harsh punishment	14
Mitigating factors that are wrong or too narrow	
Proves the defendant less likely to be guilty of first-degree murder	11
Caused the defendant to kill but did not justify or excuse	18
Excuse within the defendant or arising from defendant's background	10
Justifications arising from the circumstances of the crime	5
Described the defendant in a positive light but failed to describe the implications for responsibility or punishment	2
Suggests that the crime was less heinous	2
Mitigating factors that imply more punishment	
Suggests that the crime was more heinous	6
Proves the defendant more likely to be guilty of first degree murder	3

with either no declarative knowledge or correct knowledge are likely to easily understand the concept of "aggravation" offered in the pattern instructions. On the other hand, jurors who bring to the jury room declarative knowledge that is wrong or too narrow, as in the bottom of Table 1, will be more difficult to instruct (e.g., "Aggravating, I think is the cause or viciousness of the crime. Like the more vicious the crime is . . . like cutting people"). The pattern instructions will need to correct their initial understandings or broaden the categories of factors that they might consider as aggravating according to the specific case they are trying. Most difficult will be jurors who begin with incorrect declarative knowledge, a pre-existing belief that aggravation implies less punishment because it justifies the crime (e.g., "Aggravating circumstances would probably mean someone provoked

Table 3  
*Frequencies of Category Definitions for Aggravation and Mitigation Across Samples*

Category	n (and %) for:		
	Interview sample	Survey sample	Combined
Aggravating circumstance definitions <sup>a</sup>			
Correct/no definition	14 (19%)	14 (28%)	28 (22%)
Wrong—too narrow	19 (26%)	19 (26%)	32 (25%)
Wrong—less punishment	43 (57%)	23 (46%)	66 (52%)
Total	76 (60%)	50 (40%)	126 (100%)
Mitigating circumstance definitions <sup>b</sup>			
Correct/no definition	38 (50%)	18 (36%)	56 (44%)
Wrong—too narrow	31 (41%)	28 (56%)	59 (42%)
Wrong—more punishment	7 (9%)	4 (8%)	11 (9%)
Total	76 (60%)	50 (40%)	126 (100%)

<sup>a</sup> $\chi^2(2, N = 126) = 1.92, p > .05$ . <sup>b</sup> $\chi^2(2, N = 126) = 2.39, p > .05$ .

someone to do something—a physical threat, and extreme measure”). To be successful, pattern instructions will need to replace an incorrect knowledge structure with one that contains content directly contradicting the juror’s initial, and incorrect, understanding.

In Table 2, we used a similar system for categorizing initial declarative knowledge that jurors possess regarding mitigation. Again, some bring no understanding or a correct concept of “mitigation” and are likely to easily understand the definition of mitigation offered in the pattern instructions (e.g., “Mitigating circumstances are things that make the defendant less deserving of the death penalty”). Others begin with concepts that are either wrong or too narrow and will therefore miss potential mitigation that focuses on the strengths and limitations of the defendant, unless they are corrected with the pattern instructions (e.g., “A mitigating circumstance is a situation that explains why the murder came about”). Finally, others start with declarative information that is simply wrong, and it leads them to believe mitigation implies that more punishment is justified (“Psychologically thought through, like premeditated. Where you think about it beforehand and have it planned out—it’s conceived”). These jurors will be especially difficult to rehabilitate because the instructions will need to replace an incorrect knowledge structure with one that contains content that directly contradicts the juror’s initial, but once again incorrect, understanding.

### *Discussion*

Table 3 shows that across 126 interviewees, 77% of the respondents demonstrated declarative knowledge about aggravation that was incorrect (25% offered definitions that were too narrow and 52% thought that aggravation implied less punishment). Further, 52% showed an incorrect definition of mitigation (42% had a definition that was too narrow and 9% thought mitigation required more punishment). It is clear from these data that it is wrong to think of jurors as blank slates upon which the court engraves its recipe for guided discretion. Instead, these jurors are active processors of social information who possess initial levels of declarative knowledge that may compromise their ability to adopt the judges’ instructions. These data challenge the presumption made by the Supreme Court in *Richardson* (1987) that the jury understands the instructions as the judge provides them. Instead, the results help explain the level of confusion that is continually demonstrated in the psychological investigations of jury comprehension of penalty phase instructions in capital murder trials. Jurors bring to the judgment task a variety of misconceptions and errors concerning aggravation and mitigation that may interfere with their ability to comprehend and apply the jury instructions that the judge presents at trial. The errors lend support to the perspective that Diamond took in 1993, when she argued that that jurors’ cognitive schemata about crime and punishment may limit their ability to understand death penalty instructions. If schemata and other attributes of the cognitive system limit attention, perception, and recall (see Kunda, 1999), then it is possible that even the best and clearest instructions presented in simplified language by themselves will not adequately reduce juror miscomprehension and its impact on first-degree murder penalty decisions. Diamond concludes that “death penalty decisions may simply face a consistency ceiling that cannot be penetrated” (p. 432).

The conclusions of our work up until this point suggest the following: (a) Death-qualified citizens eligible to sit on capital murder trials miscomprehend jury instructions in the penalty phase of trials (Wiener et al., 1995); (b) although research participants make errors of understanding in both declarative and procedural knowledge that make up patterned jury instructions, they have the most difficult time with following the procedures and rules of law (Wiener et al., 1998); (c) there are specific declarative misconceptions that citizens bring to the jury task, which may explain some of the low levels of comprehension reported in the literature (i.e., the study reported in this article); and finally, (d) in two studies, mock jurors who show the greatest amount of misunderstanding were more likely to impose a death sentence than those who understood the instructions better (Wiener et al., 1995, 1998). Further, Lynch and Haney (2000) showed that miscomprehension might translate directly into racial bias imposed against African Americans (and possibly other defendants who are not of the same ethnic background as the evaluators). Next, we report the results of an elaborate and extensive jury simulation project conducted in our psychology and law research laboratory designed to extend our findings with more realistic materials that we presented to deliberating mock juries. We set out to examine some methods to improve comprehension of declarative knowledge (i.e., simplifying jury instructions and debunking common misconceptions about the law) and procedural knowledge (simplifying instructions, presenting information in a flowchart format, and offering opportunities for jurors to practice applying instructions).

### *Jury Simulation Study 1*

#### *Overview*

First, we showed to groups of research participants (death-qualified, eligible jurors) a video reenactment of a death penalty trial, which included both the guilt and penalty phases. We shortened and modified the transcripts in the Missouri first-degree murder trial, *State v. Knese* (1999), and created a 2.5-hr reenactment of the trial. Following the direct and cross examinations of witnesses, mock jurors listened to and watched a judge read jury instructions for both the verdict and sentencing decisions charged to the jury. At the end of the guilt phase, just before the beginning of the penalty phase, mock jurors learned that the defendant in the case was found guilty of first-degree murder. We modified the sentencing instructions given at the conclusion of the penalty phase, creating eight different conditions in the mock jury study. Participants completed a jury comprehension survey measuring their understanding of the death penalty instructions that they were to use to reach their sentences; they deliberated as a jury and assigned a penalty to the defendant (i.e., life in prison without the possibility of probation or parole or the death penalty). Finally, participants completed the jury comprehension survey a second time.

We presented mock jurors with one of five different sets of penalty phase jury instructions.<sup>5</sup> One set of instructions consisted of the MAI used in the Knese case.

---

<sup>5</sup>Copies of all materials, including the instructions used in each condition are available from Richard L. Wiener.

In the MAI condition, the judge in the tape read the actual instructions that were used in the trial. These instructions provide directions regarding the standard of proof (i.e., beyond a reasonable doubt), the burden of proof (i.e., the state has the burden of persuasion), and the unanimity rule (i.e., all jurors must agree upon at least one statutory aggravating circumstance and on a final punishment). They include descriptions of the role of statutory and nonstatutory aggravating and mitigating circumstances (i.e., the jurors do not need to agree upon the same mitigating circumstances and they have wide discretion in what constitutes mitigation) and a clarification of life in prison, that is, a defendant assigned to life in prison will not be paroled. The instructions describe how Missouri law requires them to weigh aggravating and mitigating circumstances, the nature of unanimous decisions, and that it is the ultimate responsibility of the jury to assign a sentence.

The second set of instructions was the control, or baseline, instructions. In the videotaped trial, the judge read the penalty phase instructions, removing descriptions of aggravating and mitigating circumstances. The judge did not define the "beyond a reasonable doubt" standard of evidence or give guidance about the responsibilities of the court and the jury in assigning a final sentence. Participants saw and heard nothing about the legal meaning of the phrase "life in prison without the possibility of probation or parole," nor did they receive direction regarding how they were to compare the aggravating and mitigating circumstances. This condition was included to control for information and misinformation that eligible jurors bring to the sentencing task (see Wiener et al., 1995). Participants instructed with the control instructions had only their own knowledge of the law to use to determine whether to invoke a death sentence or life in prison.

The third set of instructions were simplified, or rewritten, jury instructions for which the videotaped judge read MAI penalty phase instructions modified in accordance with the findings of recent research in psycholinguistics and jury decision making. These changes clarified and simplified the language so that a layperson could more easily comprehend the directions. The rewritten instructions consisted of an overview of the penalty imposition task followed by a detailed explanation of each phase of the judgment process. The instructions were presented in clearly labeled subsections. The rewritten instructions used simple language and relied on abstract legal terms only when those terms were indispensable to the meaning of the instructions. Common language definitions preceded any use of complicated legal terms. Simplified grammar avoided difficult-to-understand constructions such as double negatives and imbedded clauses. These instructions were a modified version of the model instructions used by Wiener et al. (1995). Table 4 presents examples of instructions regarding mitigation in its actual form, the baseline form, and the simplified form to illustrate the manner in which we modified the instructions. Both the MAI and the simplified instruction examples describe the unanimity rule applied to aggravation (i.e., all jurors must agree upon the same statutory aggravator to impose a death sentence) and the *Mills* (1988) mitigation rule (i.e., not all jurors must agree on the same mitigating factors in order to consider those factors in mitigation). The baseline condition contains neither of those rules; therefore, the question becomes, Do jurors presented with the simplified and MAI instructions know more declarative and/or procedural information about these rules than do jurors presented only with

(text continues on page 541)

Table 4

*Mitigating Circumstance Instructions Via the Missouri Approved Instructions (MAI), Baseline, and Simplified Instructions*


---

 MAI (actual instructions in the case)
**Instruction No. 21**

As to Count 1, if you unanimously find that the facts and circumstances in aggravation of punishment, taken as a whole, warrant the imposition of a sentence of death upon the defendant, you must then determine whether there are facts or circumstances in mitigation of punishment, which are sufficient to outweigh the facts and circumstances in aggravation of punishment. In deciding this question, you may consider all the evidence presented in both the guilt and punishment stages of trial.

As circumstances that may be in mitigation of punishment, you shall consider:

1. Whether the defendant has no significant history of prior criminal activity.
2. Whether the murder of Karen Nelson was committed while the defendant was under the influence of extreme mental or emotional disturbance.
3. The age of the defendant at the time of the offense.

You shall also consider any other facts or circumstances, which you find from the evidence in mitigation of punishment.

It is not necessary that all jurors agree upon particular facts and circumstances in mitigation of punishment. If each juror determines that there are facts or circumstances in mitigation of punishment sufficient to outweigh the evidence in aggravation of punishment, then you must return a verdict fixing defendant's punishment at imprisonment for life by the Department of Corrections without eligibility for probation or parole.

---

Baseline instructions

**Instruction No. 21**

As to Count 1, if you find that the facts and circumstances in aggravation of punishment warrant the imposition of a sentence of death upon the defendant, you must then determine whether there are facts or circumstances in mitigation of punishment, which are sufficient to outweigh the facts and circumstances in aggravation of punishment. In deciding this question, you may consider all the evidence presented in the trial.

As circumstances that may be in mitigation of punishment, you shall consider:

1. Whether the defendant has no significant history of prior criminal activity.
2. Whether the murder of Karen Nelson was committed while the defendant was under the influence of extreme mental or emotional disturbance.
3. The age of the defendant at the time of the offense.

You shall also consider any other facts or circumstances, which you find from the evidence in mitigation of punishment.

---

Simplified instructions

**Instruction No. 21—Evaluating aggravating circumstances**

. . . . If one of the jurors has reasonable doubt that the aggravating factors justify a death sentence, then the jury must sentence the defendant to life in prison without the possibility of probation or parole. However, if all jurors agree beyond a reasonable doubt that the aggravating factors justify a death sentence, then the jury may go on to consider the mitigating factors in Instruction No. 22.

Table 4 (*continued*)**Instruction No. 22—Evaluating mitigating circumstances**

You must determine whether there are mitigating circumstances present which outweigh the aggravating circumstances present. Mitigating circumstances are any facts about the crime or the defendant that make Randall Nelson less deserving of the death penalty.

Some mitigating circumstances that you must consider are:

- (a) that Mr. Nelson has no significant history of prior criminal activity,
- (b) that Mr. Nelson was under the influence of extreme mental or emotion disturbance at the time of the murder, and
- (c) Mr. Nelson's age at the time of the offense.

If any juror finds one or more of these circumstances, then he or she may consider it or them when weighing the mitigating circumstances against the aggravating circumstances. There may be other mitigating circumstances that you find in the case. You may consider any circumstance that you consider to be mitigating. It need not be named in this instruction. In deciding if a mitigating circumstance or circumstances exist, you may consider all the evidence that you heard during the guilt and penalty parts of the trial. It is not necessary for all jurors to agree that a particular mitigating circumstance exists and it is not necessary for the defense to show that it exists beyond a reasonable doubt. For a mitigating circumstance to be considered it is enough for one juror to find a circumstance, that makes Mr. Nelson less deserving of the death penalty. Although all jurors must agree beyond a reasonable doubt when determining aggravating circumstances, they do not need to do so when determining mitigating circumstances.

If each juror finds that a mitigating circumstance or circumstances exists that outweighs the aggravating circumstance or circumstances in this case, then you must sentence Mr. Nelson to life imprisonment without eligibility for probation or parole. The mitigating circumstance or circumstances that one juror finds need not be the same as those that the other jurors find.

the baseline instructions? Table 4 illustrates one example of the differences between the baseline, MAI, and simplified instructions. The full set of instructions (available from Richard L. Wiener) manipulates the presence and clarity of multiple types of declarative and procedural information.

The fourth set of instructions was based on our flowchart of the MAIs (see Figure 1). We created a graphic presentation of the jury instructions to present to participants. First, the judge read instructions regarding the beyond-a-reasonable-doubt standard, unanimous judgment, aggravating circumstances, mitigating circumstances, ultimate sentencing responsibility, and life without the possibility of probation or parole. These instructions came directly from the MAIs, as did instructions that explained how to weigh aggravating and mitigating circumstances. Supplementing the MAIs as presented in the actual case was a graphic display (in the video-tape and in an instruction booklet) of the penalty phase decision-making process. The flowchart depicted each of the decision points listed below in a manner similar to Figure 1. The judge explained how the law requires the jury to reach a decision including each of the elements needed to make a judgment. The steps are summarized below:

- Step 1.* Do all 12 jurors find at least one aggravating circumstance beyond a reasonable doubt? If the answer to this question is *no*, then the defendant must be sentenced to life in prison without the possibility of probation or parole. If the answer to this question is *yes*, then go on to Step 2.
- Step 2.* Do all 12 jurors believe beyond a reasonable doubt that aggravating circumstance(s) are sufficient to warrant the death penalty? If the answer to this question is *no*, then the defendant must be sentenced to life in prison without the possibility of probation or parole. If the answer to this question is *yes*, then go on to Step 3.
- Step 3.* Do all 12 jurors agree that the mitigating circumstances outweigh the aggravating circumstances? In answering this question, you may consider the mitigating circumstances, which were given to you in Instruction 21 [which include statutory and non-statutory mitigation]. If the answer to this question is *yes*, then the defendant must be sentenced to life in prison without the possibility of probation or parole. If the answer to this question is *no*, then go on to Step 4.
- Step 4.* Given all the circumstances in the case, do all 12 jurors agree the defendant should be put to death? If the answer to this question is *yes*, you must sentence the defendant to death. If the answer to this question is *no*, go on to step 5.
- Step 5.* Given all the circumstances in the case, do all 12 jurors agree that the defendant must be imprisoned for life without probation of parole? If the answer to this question is *yes*, then the defendant is sentenced to life in prison. If the jurors cannot agree then the Court decides the sentence.

The fifth set of instructions was debunking instructions that corrected common misperceptions about the law (hereafter, “debunking instructions”). These instructions added to the MAIs used in the case a series of comments that explained common misconceptions that people hold about the law. The debunking instructions warned the participants against making these errors. Table 5 lists some excerpts from the debunking instructions for aggravation and mitigation. As illustrated in Table 5, these instructions informed participants about the results of our interview study and warned them not to fall victim to common misperceptions about aggravation and mitigation. In addition, the instructions presented admonishments concerning the standard of evidence and ultimate jury responsibility both of which came from the results of our interview study and the work of other researchers (cf. Haney & Lynch, 1994; Haney et al., 1994). For example, with regard to the definition of *beyond a reasonable doubt*, the judge in the videotape added the following comments to the definition provided in the MAI instruction (i.e., “firmly convinced”):

Research has shown that many people hold incorrect beliefs about the definition of beyond a reasonable doubt. 1) Some people believe that to hold any proposition

beyond a reasonable doubt means that there can be *no doubt at all* about that proposition. This idea about beyond a reasonable doubt is simply wrong and should not be used in making your sentencing decision.

Further, the judge added the following admonishment to the direction that it is the primary duty and responsibility of the jury to fix the punishment:

Research suggests that many people find this confusing in that they incorrectly conclude that the judge has the final responsibility for assigning the sentence. In fact, under the law, the jury has the final sentencing responsibility.

In addition to these five variations, we asked mock jurors in three practice conditions to rehearse their application of the MAI instructions in a dummy case. Before they watched the videotape of the reenacted trial, the practice participants read a fact pattern summarizing a separate case (*Butler*, 1990) in which the defendant was convicted of first-degree murder. In one practice condition, the jury sentenced the defendant to life in prison without the possibility of probation or parole (hereinafter, "life outcome"); in a second practice condition, the jury sentenced the defendant to death (hereinafter, "death outcome"); and in the final condition, the jury's sentence was not revealed to the practice jurors (hereinafter, "no outcome"). In each condition, participants received a copy of the MAI for the penalty phase in the actual case of *State v. Butler*.

We presented the varied instructions to different mock jury panels in order to accomplish the following goals: (a) measure the comprehensibility of penalty phase jury instructions; (b) establish the amount of improvement in comprehension that can be expected for jury instructions written in less technical, more "user friendly" language; (c) improve juror comprehension, using a flowchart presentation of the decision points that make up the penalty phase instructions; (d) improve juror comprehension by correcting or debunking the schematic misperceptions and biases that jurors bring to the sentencing task; and (e) improve comprehension by allowing participants the opportunity to practice the instructions prior to using them in the target case. More specifically we hypothesized that (a) compared with the baseline and actual MAI instructions, the simplified MAI instructions would produce greater comprehension; (b) the practice, flowchart, and debunking conditions would produce greater comprehension relative to the control and the original MAI instructions; (c) all three practice conditions (life, death, and no outcome) would show greater comprehension than the baseline and original MAI conditions. However, we anticipated no differences in comprehension among the practice conditions because each requires an equivalent amount of practice in applying the law. (Note that we included the three different outcomes to test the effects of different sentence outcomes on jury sentences. We expected that the practice condition, without a sentence, would produce the same increases in comprehension as the life or death outcome conditions but that it would be less biasing than the other conditions with regard to the sentence in the target case [i.e., the Knese sentence that our participants reached]. The policy implication is that including a no-sentence practice session before the actual case would improve comprehension and not alter the sentence outcome.)

(text continues on page 545)

Table 5  
*Debunking Instructions Presented Along With the Missouri Approved Instructions*

---

Aggravating circumstances

Research has shown that many people hold incorrect beliefs about what aggravating circumstances actually are. For example:

- 1) Some people believe that an aggravating circumstance is something that caused the murderer to kill the victim. This is an incorrect idea about aggravating circumstances and should not be used in deciding your sentence for the defendant.
- 2) Other people believe that an aggravating circumstance is a factor that aggravated or provoked the defendant to kill the victim. This definition is based on the common use of the word aggravation. However, this is an incorrect definition of aggravating circumstance and should not be used in imposing a sentence upon the defendant.
- 3) Some people believe that aggravating circumstances are factors, which make the defendant more likely to be guilty of capital murder. It is true that some of the facts about the crime or the defendant that make the defendant more deserving of the death penalty may also support a guilty verdict. However, the guilt of the defendant has already been established through evidence and testimony presented during the guilt phase of the trial. Therefore, your confidence in the defendant's guilt should not be used in making a sentencing decision.

Research shows that some people are more accurate in their beliefs about aggravating circumstances, but they use definitions that are not exact. Examples of definitions of aggravating circumstances which are not exact include:

- 1) The defendant was malicious
- 2) The crime was heinous

While these factors **may** be components of aggravating circumstances, they do not offer exact definitions.

You must consider the factors, which will be defined for you as aggravating circumstances in the jury instructions when making your sentencing decision. . . .

---

Mitigating circumstances

Research has shown that many people hold incorrect beliefs about what mitigating circumstances actually are. For example:

- 1) Some people believe that a mitigating circumstance is a circumstance that makes the crime more heinous. This idea about mitigating circumstances is the opposite of the actual meaning of mitigation and should not be used in making your sentencing decision.
- 2) Some people believe that mitigating circumstances are factors which make a defendant less likely to be guilty of murder and other people believe mitigating circumstances make a defendant more likely to be guilty of murder. The guilt of the defendant has already been established through evidence and testimony presented in the guilt phase of the trial. Therefore, your uncertainty about the defendant's guilt should not be used in making a sentencing decision.

Table 5 (continued)

---

*Mitigating circumstances (continued)*

Research shows that some people are more accurate in their beliefs about mitigating circumstances, but they use definitions that exclude some mitigating factors. Although people may use one of the following definitions of mitigating circumstances, a true definition could include all of these components. Definitions of mitigating circumstances which are too narrow include:

- 1) Circumstances that caused the murderer to kill the victim.
- 2) Factors about the defendant or the defendant's background that excuse the murder.
- 3) Circumstances surrounding the crime that justify the murder.
- 4) Characteristics of the defendant that present the defendant in a positive light.

While all of these factors **are** mitigating circumstances, other factors may also be mitigating circumstances. You may consider mitigating factors to be any circumstance about the crime or the defendant, which makes the defendant less deserving of the death penalty.

---

*Method*

*Research participants.* We placed advertisements inviting participation in several local newspapers and distributed flyers to potential participants at supermarkets, churches, and recreational facilities in the greater metropolitan area of St. Louis, Missouri. Research assistants screened individuals who responded to the advertisements over the telephone to ascertain whether they met inclusion requirements. In accordance with the requirements for jury service in the state of Missouri, eligible participants were residents of Missouri, were 21 years old or older, and had no felony convictions. In addition, we invited only registered voters and/or those who possessed a valid Missouri driver's license to participate in the mock jury experiment. These additional criteria assured that our sample was representative of citizens who serve on Missouri juries because Missouri courts generate venire panel names from lists of licensed drivers and registered voters.

In addition, participants were indexed as "death-qualified" according to their responses to both the *Wainwright v. Witt* (1985) and *Witherspoon v. Illinois* (1968) standards. Although the *Wainwright* standard is controlling law, we decided to disallow participants who responded to *Witherspoon* inquires by stating that they were not able to invoke the death penalty under any conditions or that they would automatically invoke the death penalty in every case. Simulating the outcome of an actual *voir dire*, we invited only participants who stated that they could sentence a defendant to death if the circumstances of the case warranted that sentence, that they would not automatically sentence a defendant to death, and that their sentiments about the death penalty were not so strong as to prevent or substantially impair their performance in accordance with their oath as a juror.

Taking into consideration each participant's availability, researchers randomly assigned up to 14 participants to one of several weekend or evening

sessions that participants initially offered as possible testing times. We scheduled 14 participants to each session in order to compensate for those respondents who would fail to attend agreed-upon sessions. All participating individuals were informed that the study would last between 4.0 and 4.5 hr and that they would be given a \$25 stipend for participating. Each of 80 mock jury sessions was randomly assigned to one of the eight experimental instruction conditions.

Overall, 726 participants completed the study within a 1-year period (May 1999 to February 2000). Upon arrival to the mock jury room, each participant completed, among other forms an attitude toward the death penalty survey, which in counterbalanced order asked a *Witt* death-qualifying question and a *Witherspoon* (1968) death-qualifying question. Although we labeled as “death qualified” all participants at the time of the telephone screening, 14 who reported to the psychology and law laboratory agreed with the statement “I have such strong sentiments about the death penalty that *they would* seriously affect me as a juror and would prevent or substantially impair my performance in accordance with my instructions and oath.” Data from these participants were not included in the analyses reported below.<sup>6</sup> Four additional excluded respondents claimed that they would not be impaired by their attitudes toward the death penalty but nevertheless endorsed the *Witherspoon* statement

If the defendant was found guilty of a murder for which the Missouri law allowed a death sentence, I would always vote to sentence the defendant to death even if the facts in the case did not show that the defendant deserved a death sentence.

Finally, another 4 excluded respondents claimed that they could be impartial but endorsed the statement

I have such strong doubts about the death penalty that I would be unable to find the defendant guilty and vote for a death sentence where the law allowed it, even if the facts of the case showed that the defendant was guilty and deserved a death sentence.

Each of the participants who came to the laboratory for a single jury session was a member of the mock jury group for that session. In order to consider the data from any jury, we required at least 6 participants per jury.<sup>7</sup> Unfortunately, when we dropped the 22 participants who were not death qualified, 10 of 90 juries did not reach our 6-member criterion. When a panel with fewer than 6 individuals arrived at the laboratory, we completed the jury task with the individuals who attended but replaced the jury and the data with the very next jury that fulfilled our 6-member requirement. As a result, the data analyses included 80 deliberating juries with 665 jury-eligible participants. Fifty-seven percent of these jurors were in favor of the death penalty, but they would not necessarily vote for it in every case where the law allowed it (hereinafter, *pro-death penalty*). Forty-three percent

---

<sup>6</sup>The remaining 714 participants endorsed the statement “My sentiments about the death penalty *are not so strong that they would* seriously affect me as a juror and would prevent or substantially impair my performance in accordance with my instructions and oath.” We considered these participants death qualified by the *Witt* standard.

<sup>7</sup>It should be noted that we based our 6-member cutoff on the case *Ballew v. Georgia* (1978), which holds that criminal juries in the United States cannot have fewer than 6 members.

had doubts about the death penalty but believed that would be able to find the defendant guilty and to vote for a death sentence where the law allowed it (hereinafter, *anti-death penalty*). The 10 panels that did not constitute valid juries were evenly distributed across the eight experimental conditions (i.e., all experimental conditions except one had at least one unusable panel, and three conditions had two unusable panels.) Therefore, our random assignment procedure was not compromised by differential attrition from experimental conditions.

Of the original 736 participants who attended one of our sessions, we analyzed data from 665 (90%). Of these 291 (44 %) were men, 366, (55%) were women, and 8 (1%) did not answer the gender question. Participants' ages ranged from 21 to 80 years old, with a mean age of 46.8 ( $SD = 14.30$ ). Although most of the sample was Caucasians of European descent (428; 64%), a substantial minority was African American (195; 29%). Eight participants did not provide any ethnic/racial identifying information (1%) and the remaining 6% were Asian American, Hispanic or Latin American, Native American, or of another ethnic/racial background. Educational background ranged broadly from less than 6th grade ( $n = 4$ ) to graduate or professional school ( $n = 72$ ). Twenty-one percent of the sample reported that they had completed high school or less than 12 years of formal education; 37% reported some college; 23% finished college; and 19% had some graduate or professional training. In short, the sample was rather well educated.

*Materials.* The stimulus trial was a videotaped reenactment of *State v. Knese*, a trial held in the state of Missouri in which the defendant was convicted of attempted rape and first-degree murder of his wife, the young mother of the defendant's infant son.<sup>8</sup> In the case, Mr. Knese came home one night high on cocaine to find his wife and young son asleep in their beds in their trailer home. Mrs. Knese had recently given birth to the boy through Cesarean section. In the middle of the night, Mr. Knese approached his wife asking to engage in sexual relations. Mrs. Knese refused in part because Mr. Knese was still high on cocaine at the time he approached his wife. A scuffle ensued during which time Mr. Knese climbed on top of his wife and removed her clothing. Mr. Knese claimed that his wife grabbed a lamp and hit him with it cutting him severely. He explained that during the fight his wife lost control and tried to kill him with objects in the room and by strangling him. Mr. Knese held his wife down and strangled her, eventually standing on her neck until she was no longer breathing. Mr. Knese ran out of the trailer and tried to jump into the car of a neighbor who was leaving the trailer park on her way to work. Mr. Knese spun off the car and received "road burn" over sections of his body. He was picked-up by a police officer evaluated for injuries and taken to the hospital for treatment. At the hospital, Mr. Knese confessed to the homicide, denying the charge of attempted rape and claiming he killed his wife in self-defense.

Medical testimony at the trial suggested that Mr. Knese had likely stood on his wife's neck for a long time (perhaps, as long as 5 min). There was no evidence

---

<sup>8</sup>Although the videotaped reenactment was based on the trial transcript and took a great deal of the testimony dialogue from that transcript, modifications were made in the interests of shortening the video and balancing the evidence.

of sexual penetration, but there was semen in Mr. Knese's undershorts. Blood evidence showed that Mr. and Mrs. Knese bled on each other. During the penalty phase, the state offered two statutory aggravating circumstances: (a) In murdering his wife, Mr. Knese showed depravity of mind or wickedness, and (b) Mr. Knese murdered his wife while he was attempting to rape her (that is, while committing a second felony). The defense claimed three statutory mitigating circumstances: (a) Mr. Knese had no significant history of prior criminal activity; (b) Mr. Knese was under the influence of extreme mental or emotion disturbance at the time of the murder; and (c) Mr. Knese was young at the time of the murder.

An outside production team videotaped the reenactment in a moot courtroom at the Saint Louis University School of Law, providing participants with a legal atmosphere that they might have encountered had they been in an actual courtroom. Professional actors played the roles of judge, attorneys, witnesses, defendant, and family members. The videotaped reenactment of the guilt phase consisted of opening arguments, direct and cross examination of witnesses, closing arguments, and the judge reading jury verdict instructions. The guilt phase lasted approximately 1.5 hr and concluded with the jury finding the defendant guilty of first-degree murder. The videotaped reenactment of the penalty phase consisted of opening arguments, direct and cross examination of witnesses, closing arguments, and the judge reading jury instructions for the sentencing the defendant. The penalty phase of the trial lasted approximately 1 hr and concluded with the judge's charge to the jury to use the law to determine the defendant's sentence.

Trial information presented to participants was held constant across all conditions. However, each jury session received one of eight different versions of penalty phase instructions (including the three different practice conditions) to use to determine a sentence for the defendant. At the conclusion of the penalty phase, the judge read one of five penalty phase instructions (i.e., baseline, MAI, simplified, flowchart, or debunking) and the researchers provided an instruction booklet to each participant. The participants were encouraged to refer to the booklet as they completed questionnaires and deliberated. Three of the four groups of juries that received the MAI instructions practiced with the penalty phase instructions before watching the videotaped reenactment of the Knese trial. As described above, we presented a practice, capital murder trial, *Butler*, in the form of a four-page vignette. The first three pages of the summary included information regarding the guilt phase of the trial and concluded with the jury finding Mr. Butler guilty of first-degree murder. The last page included a summary of the penalty phase, which concluded with one of three outcomes: The jury sentenced the defendant to death; the jury sentenced the defendant to life in prison without the possibility of parole; or the jury did not report a punishment. Practice jurors watched and listened to a videotaped presentation of the judge reading the instructions for the case after they finished reading the written vignette. Researchers also provided these judicial instructions in a printed booklet. After the participants viewed these instructions, the experimenter handed out a list of open-ended questions pertaining to the instructions and allowed the participants time to write down their answers. To answer the questions, the participants needed to review and consider the *Butler* MAI instructions. After the participants completed the questionnaire, the experimenter read the answers to the questions. The

answers to the open-ended questions were not treated as measures and were not analyzed.

In this manner, we created eight penalty phase instruction conditions: (a) baseline instructions, (b) MAI (actual case) instructions, (c) simplified instructions, (d) flowchart instructions, (e) debunking instructions, (f) practice instructions with life in prison as the outcome, (g) practice instructions with the death penalty as the outcome, and (h) practice instructions with no stated outcome.

After viewing the entire case and again after deliberating, each mock juror completed a 36-item multiple-choice, juror comprehension survey modeled after the one used in Wiener et al. (1998). There were four types of questions. Twenty questions tested knowledge of constitutional law as decided in Eighth Amendment jurisprudence. The topics and cases that they pertained to were as follows:

1. *Reasonable doubt*: The sentencer decides whether to impose the death penalty understands and reaches a decision by the “reasonable doubt” standard of evidence (*Cage*, 1990; *In re Winship*, 1970).
2. *Jury responsibility*: The sentencer understands that the ultimate responsibility for imposing a penalty upon the defendant resides with the jury and not the court (*Caldwell v. Mississippi*, 1985).
3. *Mitigation agreement*: The sentencer may consider mitigating circumstances not agreed upon by all jurors (*Mills*, 1988).
4. *Mitigation content*: The sentencer may consider as mitigating circumstances any aspect of the defendant’s character or record or any circumstance of the offense (*Lockett*, 1978).
5. *Parole ineligibility*: If future dangerousness is at issue in the sentencing phase, the sentencer understands truthful information about parole ineligibility (i.e., in Missouri a life sentence in a first-degree murder case is without the possibility of probation or parole) (*Simmons v. South Carolina*, 1994).

An additional 16 questions tested knowledge of state law. Specifically, these questions pertained to the concepts and procedures in the MAI concerned with burden of proof, the unanimity rule, aggravating circumstances and mitigating circumstances, balancing aggravating and mitigating circumstances, and the verdict ballot.

The questions were in two formats. Half of the questions included no case facts and merely asked about semantic information that constitutes a concept in the penalty phase instructions. These items measured declarative knowledge about the law. Each item concluded with a summary statement about the law followed with a scale: *yes* (true), *no* (false), or *do not know*. Each item had one correct response. Consider, for example, the following item that measures the mock jurors understanding of the unanimity rule, that is, all jurors must agree on at least one aggravating circumstance (the same one) to impose the death penalty. This is an example of a question that pertains to the Missouri death penalty scheme as directly specified in the MAI. (The underlined answer is correct.):

After reviewing all the evidence in the case, all jurors but one were convinced beyond a reasonable doubt that a specific aggravating circumstance was true. According to the law, must the penalty be life in prison without probation or parole?

- a. yes      b. no      c. do not know

The other half of the questions presented case facts, explained how the jury applied the law to those facts, and then asked whether the jury had followed the rules specified in the law. These items measured procedural knowledge about the law. Each item concludes with a summary statement about the rule followed with a scale: *yes* (true), *no* (false), or *do not know*. Each item has one correct response. Consider, for example, the following item that measures the mock jurors' understanding of the mitigation content (i.e., an issue of constitutional law, which states that any factor about the crime or defendant can be mitigating, not simply those factors offered in the jury instructions) (the underlined answer is correct):

Only one juror believed that Mr. Nelson's testimony indicated that he was remorseful about killing his wife. This juror believed this mitigating factor outweighed the aggravating circumstances found beyond a reasonable doubt. The juror voted for life in prison. Has the juror followed the judge's instructions?

- a. yes      b. no      c. do not know

In this manner, the jury survey consisted of 20 questions testing five independent areas of constitutional law, that is, 2 declarative items and 2 procedural items for each of the 5 constitutional topics outlined above. It also included 16 questions testing separate items of state law (8 measuring declarative knowledge and 8 measuring procedural knowledge). We created two random orders of the thirty-six items that made up the knowledge component of the juror comprehension survey.

*Procedure.* After completing an informed consent statement, each participant received a survey packet containing several questionnaires. The first questionnaire in the packet included a second issuance (i.e., the "Attitude Toward the Death Penalty Questionnaire") of the *Witt* and *Witherspoon* (1968) death-qualification standards.

In the practice conditions, participants next read the summary of *State v. Butler*. When all jurors had finished reading the summary, they viewed a videotaped presentation of a judge reading the jury instructions for that trial. At the conclusion of the tape, the researchers presented the participants with a copy of the written jury instructions that the judge had just read aloud. The jurors completed a survey regarding the facts presented during the practice trial as well as questions about the legal instructions presented during the case. Following the completion of this survey, the researcher orally related the correct answers to the practice trial survey. Jurors then completed the "Jury Sentence Survey—*State v. Butler*," which assessed their recall of the defendant's sentence as presented in the vignette. Jurors completed a second juror sentence survey for the *Butler* case at the conclusion of the *State v. Knese* video presentation.

In all conditions (practice and nonpractice), researchers informed participants that they would be watching a videotaped reenactment of the first-degree murder trial *State v. Nelson*. (Note: the names and places in the Knese trial were changed

to disguise the case, which actually took place outside Kansas city, from any mock jurors who might have come across it in the media.) Researchers verbally informed participants of the differences between the guilt and penalty phases of a capital murder trial as presented under Missouri law and invited them to take on the role of a juror actually impaneled for the case. To help the jurors take the case seriously, the researchers informed them that the results of the experiment would be part of actual litigation (which they have been). The experimenter started the stimulus trial and jurors watched the guilt phase. All participants then completed a "Guilt Phase Survey" in which they were asked to either agree or disagree with the jury's verdict, guilty of first-degree murder.

Participants next watched the penalty phase for *State v. Nelson*. Following the closing arguments, the judge read the trial instructions that jurors were to follow in determining whether the defendant should be given the death penalty or should be sentenced to life in prison without the possibility of parole. These instructions varied in format depending on which of the eight experimental conditions the jury session had been assigned. After the judge had completed his verbal instructions, jurors received booklets containing the penalty phase instructions corresponding to the assigned condition. The researcher asked the participants to review the instructions and gave them some time to do so. In pilot tests, we determined that the jurors did not pay attention to the judge reading the instructions, if we offered them to the jurors while the tape was running.

Participants completed the first of two "Juror Surveys," which included questions regarding each juror's procedural and declarative knowledge of state and constitutional laws. The researchers invited the jurors to refer back to their instruction booklets to aid them in answering these questions. All juries then chose a foreperson and were given 20 min to deliberate as to whether Mr. Nelson should receive the death penalty or life in prison without the possibility of parole. At the conclusion of the deliberation period, the foreperson filled out a "Verdict Form," assigning to the defendant either a sentence of death or life in prison without the possibility of parole. The researcher recorded on videotape all deliberations for subsequent content analysis. Jurors completed a second version of the Juror Survey following the conclusions of deliberations. The first and second surveys, counterbalanced across participants, contained identical questions but were varied in the order in which the questions were presented.

Next, participants completed a "Fact Sheet," which included questions about facts presented throughout the *State v. Knese* videotape trial. Following this Fact Sheet, which was designed to determine the extent to which participants paid attention to the trial information, was a demographics questionnaire. Finally, the researcher debriefed, paid, and dismissed the participants.

## *Results*

*Overview.* The results of this study are presented in four parts. First, we present attention and manipulation check data that verify participants viewed and understood the facts in the reenacted case, *State v. Nelson*, and that they agreed with the verdict reached at the end of the guilt phase in the tape. Additional data show that participants in the practice conditions were aware of and understood the sentence in the practice case, *State v. Butler*, presented before the video reenactment of *State v. Nelson*. Second, we present data that document the reliability of

the juror survey. Third are tests of the effects of the different instruction conditions on comprehension of jury instructions. Fourth and last is a regression of type of instruction, comprehension of instructions, and attitudes toward the death penalty on the outcome of the jury sentences.

*Attention and manipulation checks.* At the end of the deliberation period, participants completed a 10-item Fact Sheet that tested their memory for and understanding of the reenacted trial. We administered the Fact Sheet at the end of the deliberation session just before the demographic survey so that it would not influence the deliberation and sentence process. The test consisted of 10 statements about the case preceded by the instruction: "Circle 'true' if the item accurately describes the video and 'false' if the item does not accurately describe the video." Following each item was a correct and incorrect answer. Examples of two items and corresponding answers are as follows: (a) The prosecution argued that the injuries on Randy Nelson's body resulted from Randy falling off a moving car (*true*), and (b) there was physical evidence in the rape kit that showed Randy Nelson had sexual intercourse with Karen Nelson (*false*). The mean item scores for the 665 mock jurors ranged from a low of 71% correct ( $SD = .46$ ) to a high of 99% correct ( $SD = .07$ ), with an overall accuracy score of 88% correct ( $SD = .11$ ) across the full set of 10 items. Thus, the mock jurors paid attention to the videotape, generally understood the facts, and were able to recognize the correct answers to almost 9 out of 10 questions about the case.

The guilt phase ended as the foreperson (a voice without a picture) announced the unanimous jury verdict, guilty of first-degree murder. At that point, the experimenter stopped the tape and presented the guilt phase survey consisting of one item, to which participants supplied a rating from 1 (*very certain that verdict was correct*) to 5 (*very certain that the verdict was incorrect*), with a midpoint of 3 (*neither certain nor uncertain that the verdict was correct*). The item read, "After reviewing all the information presented in the guilt phase of this trial, how certain are you that the jury reached the correct decision in finding Mr. Nelson guilty of first-degree murder?" The mean for the 665 mock jurors was 1.69 ( $SD = 1.11$ ). Eighty-three percent of the mock jurors indicated that they were very certain or somewhat certain that the verdict was correct. Only 10% were somewhat or very certain that the verdict was incorrect. These data show that overall, the videotape was effective in presenting a case in which the defendant was guilty of first-degree murder.

Finally, in the practice conditions (i.e., in the *State v. Butler* vignette), a manipulation check tested whether the presentation successfully conveyed to the mock jurors that the real jury had sentenced the defendant to death, sentenced the defendant to life in prison, or that there was no information about the original jury's sentence. In the condition that ended with Mr. Butler receiving a sentence of life in prison without the possibility of probation or parole, 95% of the participants accurately reported that sentence. In the conditions ending with a death sentence or no sentence, 91% and 94% of the mock jurors correctly identified those sentences, respectively. We concluded that the manipulation was successful and that the jurors in the practice conditions understood the outcome of the case. It is interesting that errors in identifying the outcome of the *Butler* case increased markedly by the end of the experimental session. When asked at the end of their deliberations and after they announced their sentence for the *Nelson*

(*Knese*) trial, only 64%, 64%, and 68% of the life-, death-, and no-sentence jurors were able to accurately report the outcome of the *Butler* trial that they had reviewed some 4.5 hr earlier. Thus, the effect of the outcome of the practice case dissipated quickly, at least in conscious memory.

*Juror survey.* Using all 713 death-qualified respondents, reliability analyses examined the consistency of mock jurors' answers to the jury comprehension survey. For the purpose of reliability analyses, we scored each item so that a correct response was scored 1.00 and an incorrect or *do not know* response was scored 0.00. Using only the items from the pre-deliberation survey (36 items), coefficient alpha was equal to .73. The number of correct responses ranged from 35 (97%) correct to 1 (3%) correct. The mean number of correct responses was 21.96 (61%) ( $SD = 5.07$ ). With only the items from the post-deliberation survey (36 items), coefficient alpha was equal to .75. The number of correct responses ranged from 35 (97%) to 0 (0%, with a mean of 22.32 [62%]) ( $SD = 5.23$ ). Combining items from both juror surveys (72 items, two replications of each item) coefficient alpha was equal to .86. The number of correct responses ranged from 67 (93%) correct to 1 (0%) correct, and the mean number of correct responses was 41.15 (61%) ( $SD = 9.74$ ). Finally, the test-retest reliability of the jury survey (correlation between the total pre- and post-deliberation correct scores) was equal to .73,  $p < .001$ . These data indicate that the jury survey exceeded current standards of internal consistency and stability reliability levels.

*Comprehensibility analyses.* For the main analyses of juror comprehension, we created an accuracy measure on the juror survey that controlled for the *do not know* answers and chance guessing: *correct responses (hits) = 1*, *incorrect responses (misses) = -1*, and *do not know responses = 0*. We calculated accuracy averages by summing the scores for all the items that constituted a subscale and dividing by the number of items that made up that scale. This procedure credited hits but did not credit *do not know* responses and penalized misses. Resulting accuracy averages could range from 1.00 (*all hits*) to -1.00 (*all misses*). An accuracy average of 0.0 could result from all *do not know* answers, an equal number of hits and misses, or some combination of hits, misses, and *do not know* answers. The mean overall accuracy percentage across all pre-deliberation items was .28 ( $SD = .25$ ), indicating that the respondents made more correct responses than incorrect responses, (an independent  $t$  test showed that this value was significantly greater than the chance value of zero,  $t[653] = 28.20$ ,  $p < .001$ ) but that scores were not very far above the 0.00 level, which respondents could have achieved by simply answering *do not know* for all the items. The mean overall accuracy percentage across all post-deliberation items was .30 ( $SD = .25$ ), once again indicating that the respondents made more correct responses than incorrect responses (an independent  $t$  test with jury group as the aggregate data point showed that this value was significantly greater than the chance value of zero,  $t[79] = 22.36$ ,  $p < .001$ ), but still not very far above the chance level score of zero.

We used a mixed model analysis of variance design to test the effects of the instructional variations on juror comprehension of the instructions. The between-groups factor was instruction type (eight levels), and there were three within-group factors: knowledge type (declarative vs. procedural), law type (state vs. constitutional issues), and deliberation (pre-deliberation vs. post-deliberation).

The dependent variable for this analysis was the accuracy score on the jury survey averaged for each of the four scale types (i.e., declarative state law, declarative constitutional law, procedural state law, and procedural constitutional law). Participant data collected before deliberation met independence assumptions of the analysis of variance because each juror acted independently of the others when completing the jury survey. However, it is not clear that juror data after deliberation are independent observations, because the jurors on each panel worked together as a group, possibly influencing one another's answers on the second jury survey administration. Although the jurors worked independently on the second juror survey, the group discussion may have influenced individual responses. On the one hand, if one is willing to make the independence assumption, the most powerful analysis for this type of data is an analysis of variance that treats each juror response as a separate degree of freedom (i.e., the aggregate data point remains the individual juror). If the assumption of independence is untenable, it is more appropriate to treat each deliberating jury as an aggregate data point and reduce the total degrees of freedom in the analysis from the full number of respondents to the number of deliberating juries. The latter approach loses power by reducing the sample size from 655 to 80; however, it gains some power back by using an aggregated jury score, thereby decreasing random error attributable to variations among individual jurors, and it does not make the assumptions of independence between jurors that may be difficult to defend. We analyzed the data using both techniques and report below only those effects that were significant with both types of models. We report the  $F$  tests for the more conservative jury aggregate results because there were no aggregate jury effects that were not significant in the individual juror analyses.

In summary, the main data aggregate for the analyses reported below is the jury. We summarized the data from 80 juries ranging in size from 6 members to 13 members (one group had 13 deliberating mock jurors). The juries had a median value of 8 members and an average of 8.31. A one-way analysis of variance using jury size as a dependent variable and the eight instruction conditions as the independent variable showed that jury size did not vary systematically across the 8 instruction conditions,  $F(7, 72) = .83, ns$ .

The 8 (instruction type)  $\times$  2 (knowledge type)  $\times$  2 (law type)  $\times$  2 (deliberation status) mixed model analysis of variance resulted in a main effects of instruction,  $F(7, 72) = 4.28, p < .001, r = .54$ ; deliberation,  $F(1, 72) = 4.38, p < .05, r = .24$ ; knowledge type,  $F(1, 72) = 137.05, p < .001, r = .81$ ; and law type,  $F(1, 72) = 110.47, p < .001, r = .78$ . The instruction type, law type, and knowledge type main effects were qualified by significant interactions of instruction by law type,  $F(7, 72) = 2.72, p < .025, r = .46$ ; knowledge by law type,  $F(1, 72) = 140.06, p < .001, r = .81$ ; and instruction by knowledge by law type,  $F(7, 72) = 3.13, p < .01, r = .48$ .<sup>9</sup> The corresponding  $r$  values show effect sizes,

---

<sup>9</sup>Using jurors as the data point, we did find a small effect size, two-way interaction between deliberation and instruction type,  $F(7, 642) = 2.05, p < .05, r = .15$ , and a small effect size, three way interaction between deliberation, instruction type, and law type  $F(7, 642) = 3.18, p < .05, r = .15$ . However, these effects were not significant in the analyses that used the jury as the aggregate, for either Deliberation  $\times$  Instructions Type,  $F(7, 72) = 1.75, ns$ , or the three way interaction of Deliberation  $\times$  Instruction Type  $\times$  Law Type,  $F(7, 72) = 1.57, ns$ . All other effects reported in the

Table 6  
*Jury Accuracy Scores Pre- and Post-Deliberation  
 and for Differences in Knowledge Type and Law  
 Type Measures Collapsed Across Instruction Type*

Factor	<i>M</i>	<i>SD</i>	<i>t</i> (79) <sup>a</sup>
Predeliberation	.28	.11	23.20
Postdeliberation	.30	.12	22.36
Declarative knowledge			
State law	.33	.15	19.55
Constitutional law	.35	.13	23.71
Procedural knowledge			
State law	.12	.14	7.56
Constitutional law	.33	.13	23.21

<sup>a</sup>All *t* values indicate that the accuracy measures are significantly greater than zero,  $p < .001$ , two-tailed tests.

which were in the moderate to large range except for the effect of deliberation, which was small.

Table 6 shows that the total mean comprehension scores pre- and post-deliberation are significantly greater than chance or guessing, as are each of the four Knowledge Type  $\times$  Law Type measures collapsed across deliberation. All *t* tests are aggregated at the jury level of analysis. The means also show that the post-deliberation scores were significantly greater than the pre-deliberation scores for overall comprehension, although the effect size was small ( $r = .24$ ). Additional *t* tests to follow-up the Knowledge Type  $\times$  Law Type interaction demonstrated that accuracy on the procedural state measure was poorer than accuracy on the declarative state,  $t(79) = 17.01$ ,  $p < .001$ ; declarative constitutional,  $t(79) = 13.91$ ,  $p < .001$ ; and procedural constitutional,  $t(79) = 15.29$ ,  $p < .001$ , measures. No other differences between conditions these within cell means were significant. These results replicate the findings in Wiener et al. (1998).

Figure 2 illustrates the main effect for instruction condition displaying the mean accuracy scores for the 36-item comprehension survey collapsed across deliberation. Tukey's post hoc tests ( $\alpha = .05$ ) showed that although the MAIs produced no better comprehension than the baseline control, the simplified instructions produced significantly greater accuracy than the baseline control and than the MAIs, whereas the flowchart and practice death outcome conditions produced significantly greater comprehension than did the baseline condition.

To explain the three-way interaction of Instruction  $\times$  Knowledge Type  $\times$  Law Type, we conducted one-way analyses of variance with Tukey's post hoc tests ( $\alpha = .05$ ) for each of the Knowledge Type  $\times$  Law Type scores collapsed

---

body of the text regarding this analysis were significant using both methods. To examine the contribution of the group interaction effect, we conducted an analysis in which jury group was nested within instruction type. After controlling for the effects of jury within instruction type, the interactions of Deliberation  $\times$  Instruction Type,  $F(7, 76) < 1.00$ , and Deliberation  $\times$  Instruction Type  $\times$  Law Type,  $F(7, 76) < 1.00$ , were not significant. We concluded that the two-way interaction Deliberation  $\times$  Knowledge Type and Deliberation  $\times$  Instruction Type  $\times$  Law Type were not reliable and analyzed them no further.

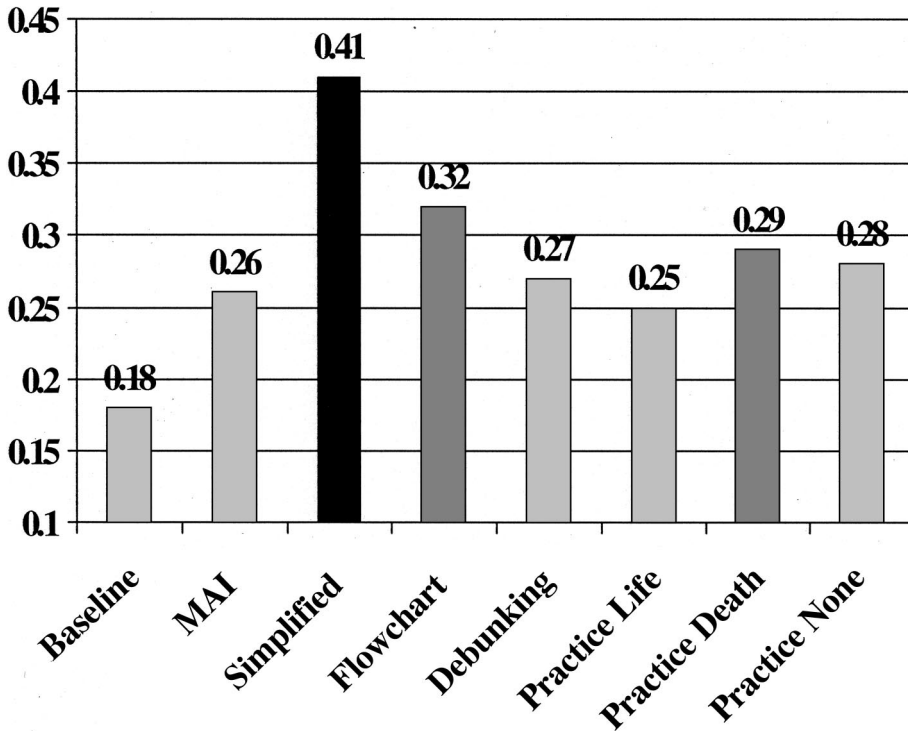


Figure 2. Percentage accuracy (hits minus misses) for total comprehension collapsed across deliberation for all instruction types. Black shading indicates values that are significantly greater than baseline and Missouri Approved Instructions (MAI); dark gray shading indicates values that are significantly greater than baseline.

across deliberation (pre- and post-deliberation combined). Figure 3 displays the results for declarative state law comprehension,  $F(7, 72) = 3.79, p < .001, r = .52$ . It shows that the simplified and flowchart conditions produce greater comprehension than the baseline condition; once again, the MAIs were not significantly different from baseline. Figure 4 shows that with regard to the declarative constitutional law accuracy effect across instructions,  $F(7, 72) = 5.06, p < .001, r = .57$ , the simplified, debunking, and practice with a death outcome conditions produced significantly greater accuracy than the baseline condition. Still, the MAI accuracy was no better than baseline for this measure. Figure 5 illustrates that simplified and flowchart instructions exceed the baseline accuracy, explaining the instruction effect on procedural state law,  $F(7, 72) = 2.68, p < .025, r = .46$ , and Figure 6 shows that the simplified instructions exceeds the baseline and the MAI instructions explaining that instruction effect on procedural constitutional law,  $F(7, 72) = 2.97, p < .01, r = .47$ .<sup>10</sup> In Figures 5 and 6, the MAI accuracy is not significantly different from the baseline condition for procedural state law or

<sup>10</sup>Note that the simplified condition produces significantly greater procedural constitutional law accuracy than do all the other conditions.

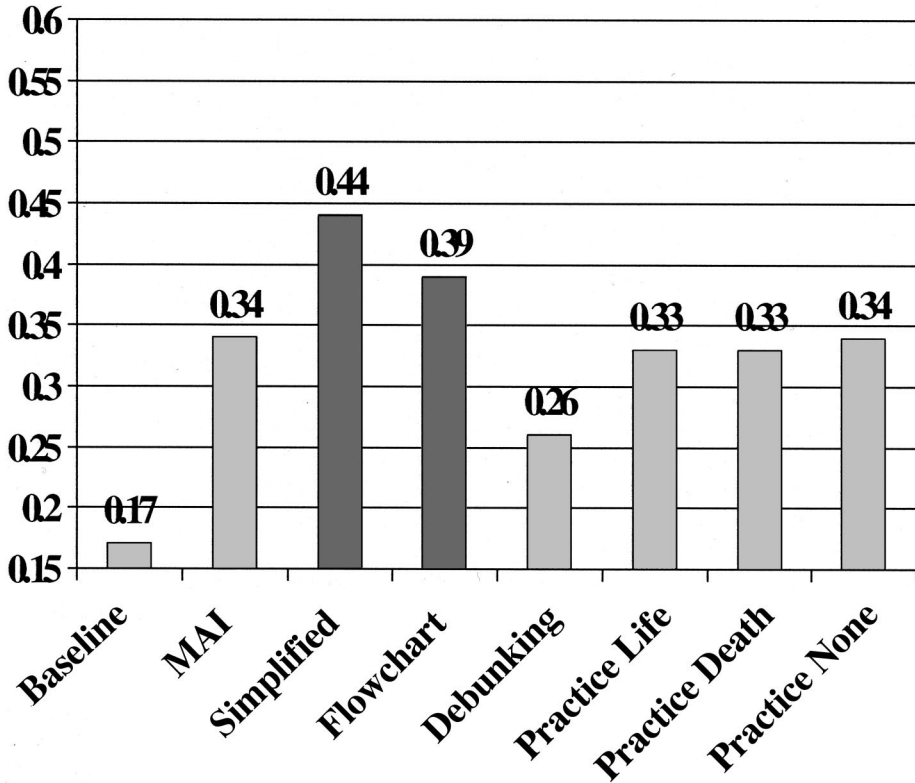


Figure 3. Percentage accuracy (hits minus misses) for declarative state law collapsed across deliberation for all instruction types. Dark gray shading indicates values that are significantly greater than baseline. MAI = Missouri Approved Instructions.

procedural constitutional law accuracy. The data displayed in Figures 2–6 demonstrate that the MAI instructions used in the actual case produced no better accuracy than the baseline condition that we stripped of critical declarative (i.e., definitions of key concepts) and procedural information (e.g., rules for weighing aggravating and mitigating factors, unanimity rules, and mitigation agreement rules). However, simplified instructions, flowchart instructions, debunking instructions (to a lesser extent), and practice (with a death outcome) increased significantly juror comprehension of at least some components of the declarative and procedural content of the law in the penalty phase.

To examine more closely jurors comprehension of the constitutional issues central to the last several decades of Eighth Amendment jurisprudence regarding the death penalty (i.e., concerns about the arbitrariness of invoking the death penalty), we conducted a separate mixed factor analysis of variance, using only the 20 items on the jury survey that measured percentage accuracy of the reasonable doubt, jury responsibility, mitigation agreement, mitigation content, and parole ineligibility issues. In the resulting 8 (instruction type)  $\times$  2 (deliber-

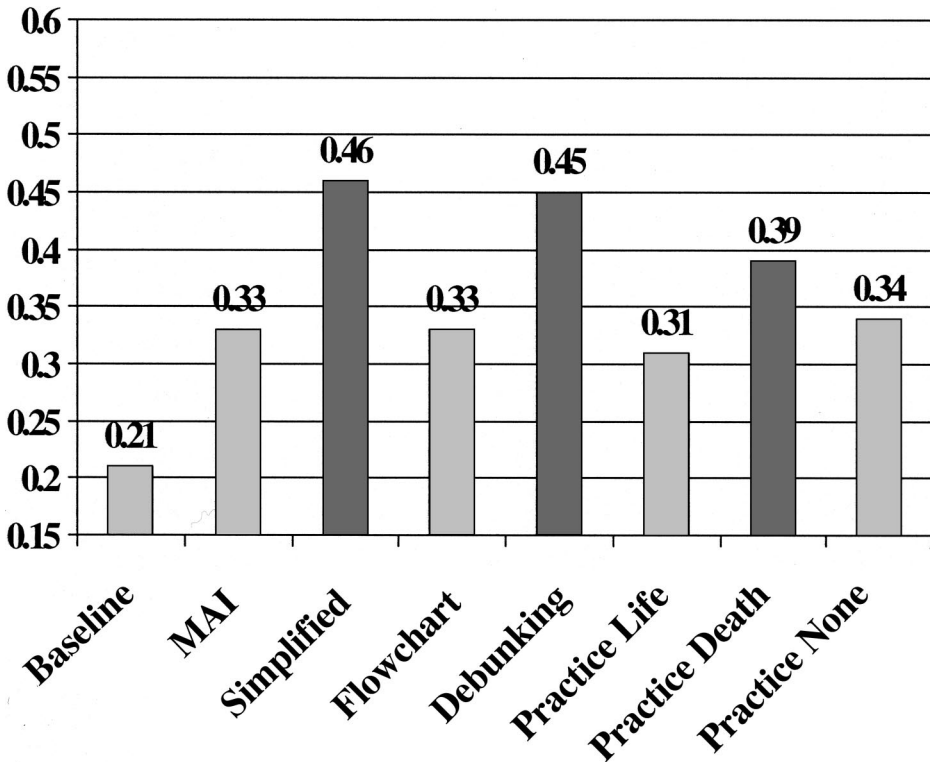


Figure 4. Percentage accuracy (hits minus misses) for declarative constitutional law collapsed across deliberation for all instruction types. Dark gray shading indicates values that are significantly greater than baseline. MAI = Missouri Approved Instructions.

ation)  $\times$  5 (constitutional issue) model with the first factor between-groups and the last two within-groups, the five measures of constitutional law issues were not separated into declarative and procedural knowledge measures. Using jury as the aggregate data point and the multivariate approach to repeated measures, we found main effects for deliberation,  $F(1, 72) = 4.08, p < .05, r = .23$ , and constitutional issue,  $F(4, 69) = 343.78, p < .001, r = .98$ . Because we analyzed the deliberation main effect with the full data in the prior analyses, we examined it no further with the constitutional measures. The very strong main effect for constitutional issue showed some startling findings. Mean accuracy percentages from lowest to highest (all significantly different from 0 at the .001 level of significance via two-tailed  $t$  tests) were reasonable doubt,  $M = -.25, t(79) = -15.26$ ; mitigation content,  $M = .30, t(79) = 10.86$ ; jury responsibility,  $M = .38, t(79) = 17.08$ ; mitigation agreement,  $M = .53, t(79) = 30.52$ ; and parole ineligibility,  $M = .75, t(79) = 33.49$ . Using the Bonferroni technique, we set alpha to .05 (actual alpha for each test was .005) and conducted all possible  $t$  tests between measures (10 tests) and found that all measures were significantly different from all other measures except that mitigation content and jury responsibility were not different from each other. These data show that across deliberation and instruction

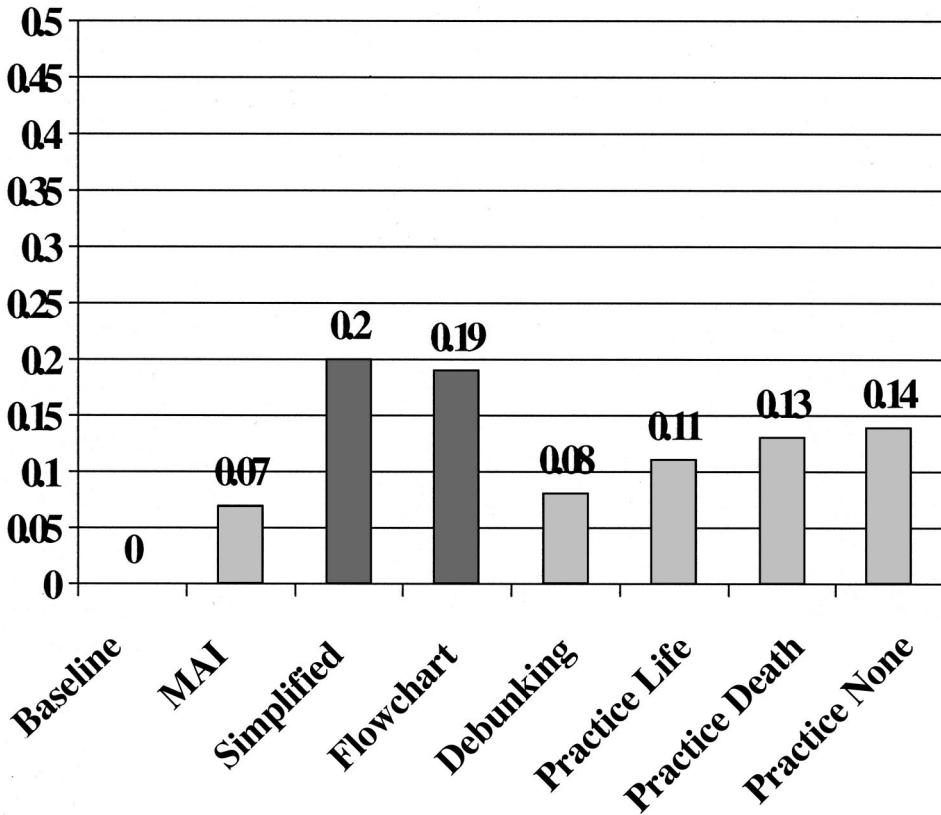


Figure 5. Percentage accuracy (hits minus misses) for procedural state law collapsed across deliberation. Dark gray shading indicates values that are significantly greater than baseline. MAI = Missouri Approved Instructions.

condition, jurors who watched our videotaped reenactment scored lower than chance on their understanding of the standard of proof, poorly on mitigation content and jury responsibility, fair on mitigation agreement, and fairly well on parole ineligibility. As in Wiener et al. (1995), mock jurors were reliably confused about the meaning of beyond a reasonable doubt in the penalty phase of capital murder trials.

The effect of scale was qualified by a Deliberation  $\times$  Constitutional Issue interaction, with (Pillai's trace) multivariate,  $F(4, 69) = 6.11, p < .001, r = .51$ , and an Instruction Type  $\times$  Constitutional Issue interaction, with (Pillai's trace) multivariate,  $F(28, 250) = 2.38, p < .001, r = .43$ . (Note that the Huynh-Feldt method of adjusting degrees of freedom for repeated measure designs produced the same effects.) Simple effects for the deliberation interaction were paired  $t$  tests comparing each constitutional measure pre- and post-deliberation. We found a deliberation effect for jury responsibility,  $t(79) = -2.14, p < .05$ , with the pre-deliberation mean for the juries equal to .40 and the post-deliberation mean equal to .36, demonstrating a slight decrease in comprehension following deliberation across all instruction conditions. A second significant effect was found for

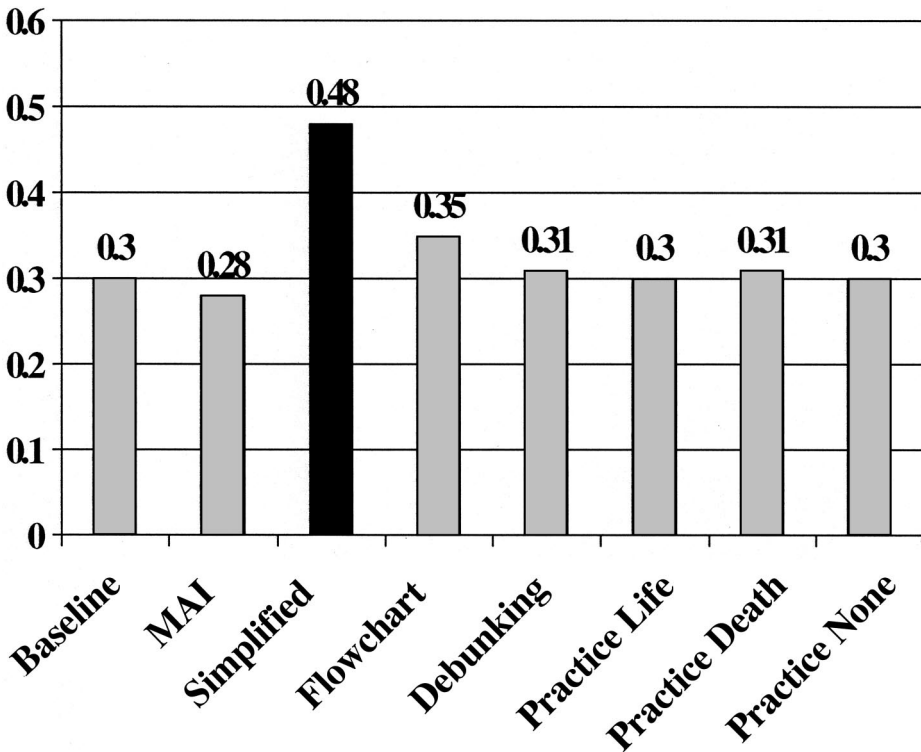


Figure 6. Percentage accuracy (hits minus misses) for procedural constitutional law collapsed across deliberation for all instruction types. Black shading indicates values that are significantly greater than baseline and Missouri Approved Instructions (MAI).

mitigation content,  $t(79) = 3.24, p < .01$ , with the pre-deliberation mean equal to .26 and the post-deliberation mean showing a slight increase in comprehension ( $M = .34$ ) across instruction type. No other deliberation simple effects were significant.

A simple effects test for the interaction of Instruction Type  $\times$  Constitutional Issue produced significant effects for mitigation content,  $F(7, 72) = 3.91, p < .001$ , and parole ineligibility,  $F(7, 172) = 5.05, p < .001$ . Figure 7 displays the results of Tukey's post hoc tests for mitigation content ( $\alpha = .05$ ), which shows that the simplified instructions produced greater accuracy than all other conditions except the debunking instructions and the practice condition with a death penalty outcome. The MAI instructions were, once again, not different from the chance baseline. Figure 8 displays the post hoc comparisons for parole ineligibility (Tukey's  $\alpha = .05$ ) and shows that all the instructional variations (except practice without a sentence) produced greater accuracy than the baseline instructions. The MAI instructions were, once again, no different from the chance baseline.<sup>11</sup>

<sup>11</sup>Using jurors as the data point, we did find an additional simple effect for condition on mitigation agreement,  $F(7, 657) = 2.30, p < .05$ . However, this effect was not significant for the

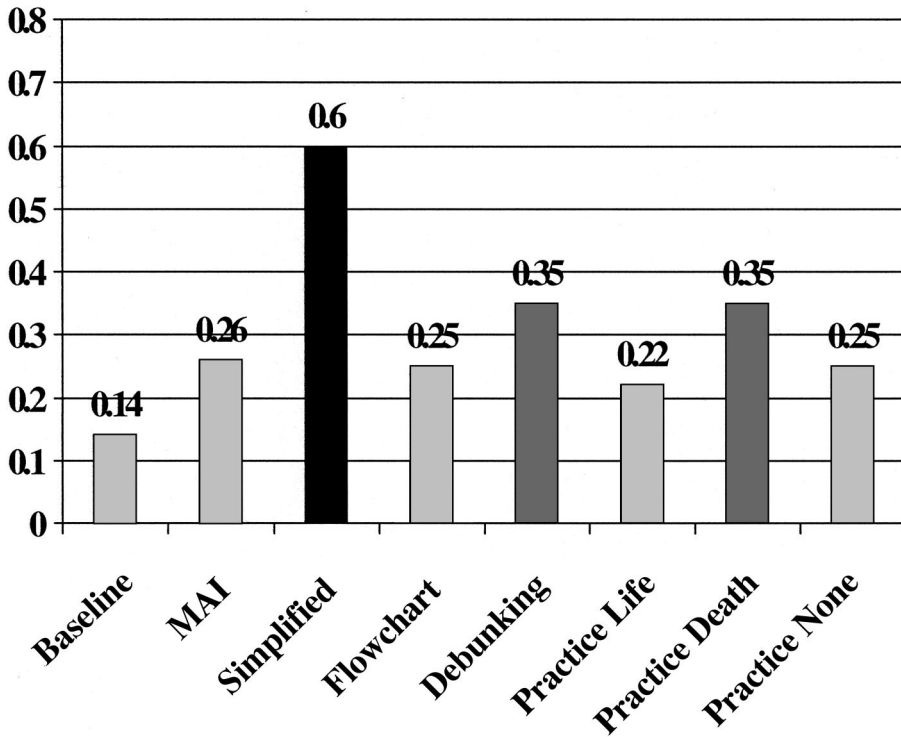


Figure 7. Percentage accuracy (hits minus misses) for mitigation content collapsed across deliberation for all instruction types. Black shading indicates values that are significantly different from those with light gray shading but not from those with dark gray shading. MAI = Missouri Approved Instructions.

We conducted a logistic regression analysis to determine the relationship between instruction accuracy and the final verdict following deliberation. Out of 80 deliberating juries, 40 unanimously agreed to assign life in prison without the possibility of probation or parole, 3 unanimously agreed to assign the death penalty, and 37 were split between the death penalty and life in prison. The criterion in the logistic regression was whether the jury found unanimously to assign life in prison to Mr. Nelson. In Missouri, juries that are unable to reach a unanimous sentence in a capital murder trial, by default, send the case to the judge for a final verdict, which may ultimately result in the death penalty. Therefore, juries can assure life in prison without probation or parole only by unanimously reaching that conclusion. For this reason, collapsing jury sentences that are not

---

analysis that used the jury as the aggregate,  $F(7, 72) = 1.43, ns$ . All other effects reported in the body of the text regarding constitutional issues were significant using both methods. To examine the contribution of the group interaction effect, we conducted an analysis in which jury group was nested within instruction type. After controlling for the effects of jury within instruction type, the effect of instruction type on mitigation agreement was not significant,  $F(7, 73) < 1.00$ . We concluded that the instruction type on mitigation agreement was not reliable in this data set and analyzed it no further.

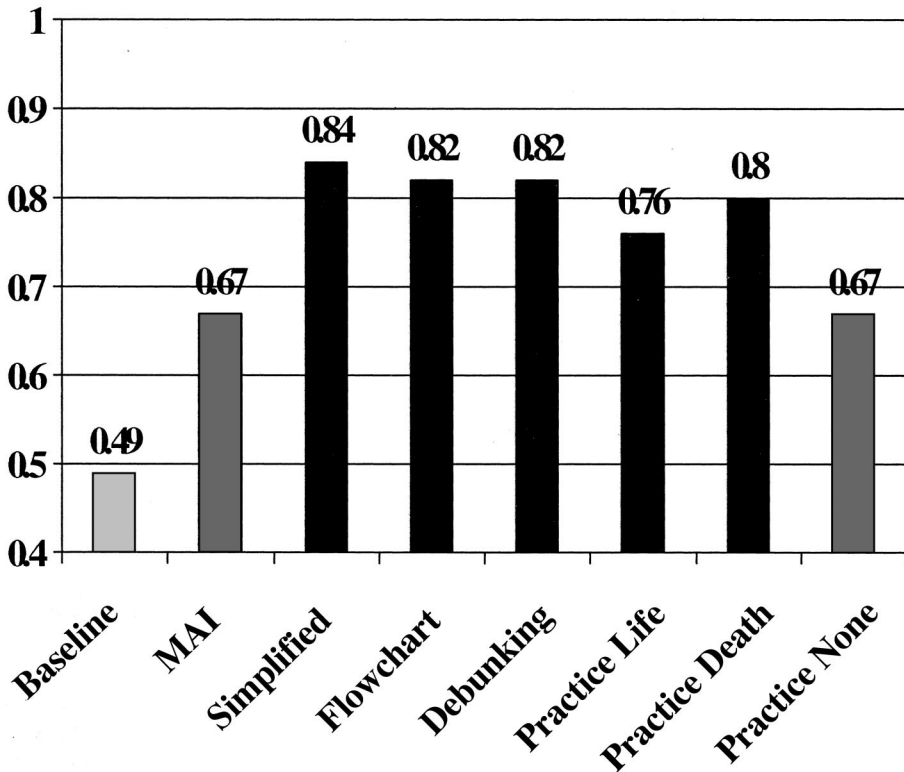


Figure 8. Percentage accuracy (hits minus misses) for parole ineligibility collapsed across deliberation for all instruction types. Black shading indicates values that are significantly different from the light gray shaded bar but not from dark gray bars. MAI = Missouri Approved Instructions.

unanimous life-in-prison sentences categorize together those cases that either assure a death sentence (i.e., unanimous impositions of death) or that risk the imposition of death from the bench (i.e., split judgments).

The predictors in the logistic regression were the four legal knowledge scales (declarative state law, declarative constitutional law, procedural state law, and procedural constitutional law), the percentage of jurors favoring the death penalty as compared with those with doubts about it (attitude toward the death penalty as determined from jurors' answers to the *Witherspoon*, 1968, death qualifying question), the numbers of jurors in each jury (a control factor), and the instruction type (i.e., the one of the eight instruction sets to which each jury was exposed). Table 7 lists the results of the analysis.

The only variable that predicted the likelihood of the jury reaching a life-in-prison sentence was the jury's aggregate accuracy score for procedural knowledge of state law. Those jurors who better knew how to carry out Missouri's death penalty scheme, that is, those who knew how to apply the rules of law to the case, were more likely to impose a life-in-prison sentence without the possibility of probation or parole than were those who scored lower on the measure. The model displayed in Figure 7 produced a significant fit to the data,  $\chi^2(1) = 5.88, p < .02$ .

Table 7  
*Logistic Regression Analysis of Likelihood of Assigning Life in Prison*

Factor	<i>B</i>	<i>df</i>	Wald $\chi^2$	<i>p</i>	Exponent of $\beta$
Procedural state law	4.24	1	5.36	.025	69.84
Declarative state law		1		.31	
Procedural constitutional law		1		.70	
Declarative constitutional law		1		.26	
Instruction type		7		.38	
% favoring death penalty		1		.17	
No. of jurors		1		.26	

It is noteworthy that neither the instructional type nor the percentage of jurors who favored the death penalty without doubts (i.e., attitude toward the death penalty) were significant predictors of deliberation outcome. Further, interpreting the exponent of the procedural state law logistic regression beta weight finds that a jury whose accuracy score was perfect 1.00 would be 69 times more likely to unanimously assign a life-in-prison sentence than would a jury with only a chance level of accuracy. Therefore, increasing the accuracy score in state procedural knowledge even with a much more modest improvement (i.e., a change smaller than moving from chance to perfect knowledge) would also be associated with a substantial increase in the likelihood of a life sentence.

### *Discussion*

The data from this mock jury study show that we reliably measured participant knowledge of penalty phase law as the videotape judge presented it in the reenactment of the *Knese* trial. Participants paid close attention to the videotape as measured by their scores on a series of factual test items. Those in the practice conditions were able to announce the outcome of the practice case and the far majority of the participants across instruction conditions agreed that the defendant was guilty of first-degree murder.

Our percentage accuracy analyses measured juror comprehension of a variety of different types of knowledge and legal constructs. In keeping with prior research, our analyses of the comprehension data pre- and post-deliberation showed that jurors' demonstrated levels of accuracy were indeed rather low. Although as a group they performed better than chance guessing with regard to overall comprehension, declarative knowledge of constitutional issues, declarative knowledge of state law, procedural knowledge of constitutional issues, and procedural knowledge of state law, on no scale did their accuracy scores exceed .35 on a -1.00 to 1.00 scale. In fact, their knowledge of procedural state law overall was only equal to .12. Although this value is greater than chance guessing, the 95% confidence interval for this measure (using jury as an aggregate) ranges from .08 to .15. This means that the percentage accuracy score on this factor in the population could be as low as .08.

When we examined the 10 mock juries (87 participants) that used the instructions from the actual *Knese* trial (not including those in the practice conditions), the accuracy percentages were .26 (declarative constitutional law),

.33 (declarative constitutional law), .28 (procedural constitutional law), and .08 (state procedural law). Although each of the first three accuracy scores is greater than 0, the state procedural law score is not. In fact, the 95% confidence interval for state procedural law using the instructions supplied at the Knese trial ranges from  $-.07$  to  $.23$ . In other words, the 87 mock jurors provided with the MAI instructions offered at Mr. Knese's trial scored no higher on state procedures than if they had simply guessed at the answers or marked all the answers with the *do not know* option.

This finding is serious enough in and of itself. It is an "empirical factor" that suggests rather strongly that jurors sentence defendants in capital murder cases with inadequate grounding in the accepted penalty phase procedure. Remember that the procedural state instructions direct the jurors on how to apply the unanimity rule to aggravation, when to look beyond aggravation to mitigation, and the manner in which Missouri statute requires the jurors to balance aggravating and mitigating factors. These data support a conclusion that the MAIs provide little or no guidance beyond chance on these important rule applications.

However, our findings regarding the impact of different instruction variations suggest the situation is even direr. First, although deliberating juries scored higher than baseline on overall accuracy in the simplified instruction, flowchart instruction, and practice conditions, they failed to do so when instructed with the MAI instructions presented exactly as they appeared in the *Knese* trial. That is, the MAI instructions offered no more knowledge than simply telling jurors to weigh aggravation against mitigation without any description or definition of key concepts (declarative knowledge) and without directions on how to apply the rules of law to the facts of the case (procedural knowledge). Further, those jurors in the simplified instruction condition outperformed the MAI instructed jurors on overall instruction accuracy. In short, the MAI instructed jurors did not understand the instructions well; however, there are several instruction modifications available that could significantly increase comprehension.

More specifically, our analyses show that simplified instructions produce increased accuracy in declarative state, declarative constitutional, procedural state, and procedural constitutional law. The flowchart approach produced relative increases in accuracy in declarative state law and most importantly in procedural state law issues. We anticipated effects of flowcharting with procedural issues because we expected that the graphic illustrations of the rules would assist jurors in understanding how the state expected them to apply the law. Like E. R. Smith (1989), we found procedural aids not only improved procedural knowledge but also increased knowledge of abstractions and concepts related to the content of the practice (i.e., declarative knowledge). We also found that the debunking instructions and practice (with a death outcome) produced increases in declarative constitutional law accuracy. Apparently, correcting errors and giving practice (with death penalty outcomes) along with simplifying the instructions allows jurors to more easily grasp the concepts that make up penalty phase law.

Finally, our analysis of sentencing across deliberating juries highlighted the importance of the findings that the jurors scored the poorest on the procedural state knowledge questions and that two methods (simplifying instructions and presenting flowchart diagrams) can increase accuracy with regard to these issues. That analysis showed that those juries who showed the greatest comprehension of

procedural state knowledge were more likely to unanimously agree on a sentence of life in prison without the possibility of probation or parole. There was a strong relationship between understanding of procedural state knowledge and a life sentence even after controlling for the type of instruction presented, the number of jurors in each deliberating jury, and the percentage of jurors favoring (as opposed to doubting) the death penalty in capital cases (i.e., attitude toward the death penalty). In fact, neither the type of instruction given nor the distribution of prior attitudes toward the death penalty influenced the sentence; only juror knowledge of the penalty phase instructions and how to apply them predicted whether juries imposed life in prison. We must qualify this finding by acknowledging that our measure of the attitude toward the death penalty was restricted to "death-qualified" jurors. That is, all mock jurors were willing and able to invoke the death penalty but were not inclined to automatically impose it regardless of the facts of the case. It is possible (indeed likely) that we would have found different results had we included jurors with more extreme attitudes toward the death penalty or if we had examined attitudes at the individual juror level rather than at the level of deliberating juries (see O'Neil et al., 2004). There is a long literature that demonstrates a relationship between juror judgments and attitudes toward the death penalty, at least at the individual juror level (see Goodman-Delahunty, Greene, & Hsiao, 1998, for a recent study and a review of that literature).

Our data also show that the simplified instructions produced greater accuracy with regard to understanding the concept of mitigation as compared with all other variations except the debunking and practice condition with a death penalty outcome. Once again, the MAI instructions were no different from the chance baseline with regard to accuracy on this essential issue. Further, with regard to knowledge about parole ineligibility all the instructional variations (except practice without a sentence) produced greater accuracy than the baseline instructions. The MAI instructions were no different from the chance baseline on this issue either.

There is one additional issue to address before concluding that existing psychological "empirical factors" do indeed show that juror miscomprehension of instructions is a serious impediment for guided discretion because of limitations in both declarative and procedural knowledge. Our data do show a small, but significant, improvement in overall comprehension post-deliberation relative to pre-deliberation across all instruction conditions. This effect is difficult to interpret because the data from this study, like many other data sets that look at deliberation effects, do not separate the effects of practice with the comprehension measure from deliberation. That is, all participants at post-deliberation had the advantage not only of deliberation but also of practice at completing the juror survey. It could be that the effect of deliberation results from the opportunity to discuss the facts, the case, and the law among jurors or it could be that the effect is largely due to the impact of practicing with the juror survey. Jury Simulation Study 1 controlled for the practice effect by using two different forms of the survey for each jury. Although both forms contained the same items, there were two different random orders of the items. Each juror answered one form at pre-deliberation and the other form at post-deliberation, with random assignment within juries with regard to which form was completed first and second. The

effects of practice were very likely diluted by using two different randomized item orders.

Nonetheless, the effects of completing each item a second time were not completely controlled. Before concluding that there is a small, but reliable, deliberation effect it is necessary to rule out practice effects as a plausible rival hypothesis. To do this, we collected data from another twenty juries.

### *Jury Simulation Study 2*

#### *Overview*

In a second jury simulation study, we collected data from another 20 juries, half of which completed both the pre- and post-deliberation Juror Surveys. The remaining jurors completed the Juror Survey after deliberations only. All juries were instructed with the MAI instructions used at the actual Knese trial. Comparing the accuracy scores pre-deliberation from one set of juries to post-deliberation from the other set of juries (the ones who did not complete a pre-deliberation jury survey) should reflect the effects of deliberation, independent of any practice effects associated with completing the survey twice. Further, comparing the accuracy scores pre- and post-deliberation for the 10 juries that completed both surveys will measure the effect of practice and deliberation in jurors instructed only with the MAIs.

#### *Method*

*Participants.* We recruited and assigned participants to two jury conditions in a manner identical to the first simulation study. Overall, 206 participants completed the second simulation study in a 3-month period (February 2000 to May 2000). Upon arrival to the mock jury room, each participant completed the attitude toward the death penalty survey, which, in counterbalanced order, asked our *Witt* death qualifying question and our *Witherspoon* (1968) death qualifying question. Although research assistants “death qualified” all participants at the time of the telephone screening, 8 who reported to the psychology and law laboratory agreed with the statement

I have such strong sentiments about the death penalty that *they would* seriously affect me as a juror and would prevent or substantially impair my performance in accordance with my instructions and oath.

Three additional participants who claimed that they could be impartial on the *Witt* question endorsed the statement

I have such strong doubts about the death penalty that I would be unable to find the defendant guilty and vote for a death sentence where the law allowed it, even if the facts of the case showed that the defendant was guilty and deserved a death sentence.

We did not include the data from these 11 participants, whom we considered not to be death qualified, nor did we include data from another participant who admitted to being a convicted felon on our demographic survey. We considered the remaining 194 participants to be death qualified and eligible to serve on a

Missouri capital murder case. In order to meet the 6-participants-per-mock-jury rule, we were forced to drop two panels and the corresponding 9 jurors from the final set of usable data. One dropped juror was from the group of jurors who completed the juror survey before and after the deliberations, and another dropped jury was from the condition that completed the survey post-deliberation only.

Of the original 206 participants who attended one of our sessions, we analyzed data from 185 (89%). Of the 184 respondents who supplied their gender on the demographics questionnaire, 77 (42 %) were men and 107, (55%) were women. Participants' ages ranged from 21 to 79 years old, with a mean age of 53.90 ( $SD = 14.77$ ). Although most of the sample were Caucasians of European descent (118; 64%), a substantial minority were African American (54; 29%). One participant did not provide any ethnic/racial identifying information (1%), and the remaining 6% were Asian American, Hispanic, Native American, or from another ethnic/racial origin. Educational background ranged broadly from less than ninth grade ( $n = 2$ ) to graduate or professional school ( $n = 46$ ). Twenty-three percent of the sample reported that they completed high school or less; 32% reported some college; 20% finished college; and 25% had at least some graduate or professional training. Overall, the sample mock jurors were slightly older but overall were very similar to the first simulated jury study sample.

*Procedure and materials.* The participants were recruited and assigned to one of two conditions, pre- and post-deliberation survey condition versus the post-deliberation survey-only condition exactly as they were in the first simulation study. The materials and case presentation were identical to those used in the MAI condition in the first simulation study except that half completed the jury survey only once, that is, at the end of deliberation. In all other ways, the procedure for this replication was identical to the MAI condition in the first simulation study.

## Results

*Overview.* We present the results of this study in four parts. First, we present attention data that verify that participants viewed and understood the facts of the case. Second, we present data that document the reliability of the juror survey; third, an analysis of accuracy statistics for the four knowledge type and law type scales; and fourth, an analysis of the effects of deliberation and practice on the jury survey for overall accuracy.

*Attention checks.* At the end of the deliberation period, participants completed the same 10-item Fact Sheet that participants completed in the first simulation study. The average item scores for the 184 mock jurors ranged from a low of 65% correct ( $SD = .48$ ) to a high of 100% correct ( $SD = .00$ ), with an overall accuracy score of 88% correct ( $SD = .11$ ). As in the first simulation, we concluded that the mock jurors paid attention to the videotape and understood the facts as presented.

At the end of the guilt phase, we again asked the mock jurors to indicate how certain they were that Mr. Nelson (Knese in the actual case) was guilty of first-degree murder on a scale from 1 (*very certain that verdict was correct*) to 5 (*very certain that the verdict was incorrect*). The mean (and standard deviation) for the 184 mock jurors was 1.68 ( $SD = 1.09$ ). Eighty-four percent of the mock jurors indicated that they were very certain or somewhat certain that the verdict

was correct. Only 10% were somewhat or very certain that the verdict was incorrect. As in Jury Simulation Study 1, the videotape was effective in presenting a case in which the defendant was guilty of first-degree murder.

*Juror Survey.* Using all 185 respondents from the 20 juries, reliability analyses demonstrated that the jury survey produced acceptable internal consistency coefficients. The reliability scoring for this study was identical to that for the first simulation study, in that correct responses were scored as 1.00 and incorrect or *do not know* responses were scored as 0.00. First, using only the items from the pre-deliberation survey (36 items) and respondents from only the pre-deliberation condition ( $n = 95$ ), coefficient alpha was equal to .70. The number of correct responses ranged from 32 (89%) correct to 9 (25%) correct. The mean number of correct responses was 21.49 (60%) ( $SD = 4.73$ ). Second, with only the items from the post-deliberation survey (36 items) and all participants, coefficient alpha was equal to .76. The number of correct responses ranged from 33 (92%) correct to 2 (6%) correct. The mean number of correct responses was 20.83 (58%) ( $SD = 5.36$ ). Third, with items from both juror surveys (72 items) and only respondents from the pre-deliberation condition ( $n = 95$ ), coefficient alpha was equal to .86. The number of correct responses ranged from 62 (86%) correct to 23 (32%) correct and the mean number of correct responses was 43.22 (60%) ( $SD = 9.40$ ). Finally, the test-retest reliability of the jury survey (correlation between the total pre- and post-deliberation correct scores,  $n = 95$ ) was equal to .73 ( $p < .01$ ). These data indicate that, as in the first simulation study, the jury survey exceeded current standards for internal consistency and stability reliability.

*Comprehensibility analyses.* For the main analyses of juror comprehension, we created the same percentage accuracy measures as in the first jury simulation study, where accuracy averages could range from 1.00 (*all hits*) to -1.00 (*all misses*). A 0.00 accuracy average could result from all *do not know* answers, an equal number of hits and misses, or some combination of hits, misses, and *do not know* answers. Table 8 presents the results for the juries who answered both pre-deliberation and post-deliberation juror surveys and for the juries who answered only the post-deliberation surveys. All values represent data points at the jury level of aggregation.

The mean overall accuracy percentage across all pre-deliberation items for the 10 jurors that completed the pre- and post-deliberation surveys was .22 for both surveys. As indicated in Table 8, both values are significantly greater than the chance value of zero, indicating that the MAI instructions produced an accuracy score that exceeded chance guessing; however, there was no gain in accuracy that could be attributed to either deliberation or the practice effect of taking the survey twice, paired  $t(9) < 1.00$ , *ns*. Further, although all the pre-/postdeliberation jurors scored significantly greater than chance on each measure, except the procedural state law measure, there were no significant differences from the pre- to post-deliberation jury survey accuracy scores on measures, as determined by paired  $t$  tests ( $\alpha = .05$ ). Replicating the effects in the first jury simulation study, the mock jurors with the MAIs scored no greater than chance (and lowest) on the procedural state law measure. Finally, as illustrated in the bottom panel of Table 8, those jurors instructed with the MAIs and who filled out the jury survey after deliberation only showed the same pattern of results with accuracy scores greater than chance on all measures except procedural state law. Although all the mean

Table 8  
*Jury Accuracy Scores for Pre- and Post-deliberation Jury Conditions*

Factor	<i>M</i>	<i>SD</i>	<i>t</i> (9)	<i>p</i>
Juries completing pre- and post-deliberation surveys:				
Predeliberation				
Total accuracy	.22	.09	8.16	< .001
Declarative state law	.28	.17	5.12	< .001
Declarative constitutional law	.30	.06	14.70	< .001
Procedural state law	-.02	.09	-.59	<i>ns</i>
Procedural constitutional law	.32	.10	9.97	< .001
Juries Completing Pre- and Postdeliberation Surveys:				
Postdeliberation				
Total accuracy	.22	.12	5.65	< .001
Declarative state law	.23	.24	3.01	< .025
Declarative constitutional law	.32	.07	15.36	< .001
Procedural state law	.03	.11	.71	<i>ns</i>
Procedural constitutional law	.31	.17	5.77	< .001
Juries completing post-deliberation surveys only:				
Post-deliberation				
Total accuracy	.17	.10	5.28	< .001
Declarative state law	.19	.19	3.11	< .025
Declarative constitutional law	.24	.08	9.26	< .001
Procedural state law	.03	.16	.63	<i>ns</i>
Procedural constitutional law	.22	.16	4.44	< .01

*Note.* The *t* values test the accuracy measures against the population value of zero (two-tailed tests).

accuracy scores on the pre-deliberation measure (for the pre- and post-deliberation groups) were greater than those on the post-deliberation measures for the post-deliberation survey only group, only the independent *t* test for the declarative constitutional law measure showed a significant effect,  $t(18) = 2.11$ ,  $p < .05$ . Ironically, the means illustrated in Table 8 suggest that when instructed with the MAIs, deliberation lowered the jurors' accuracy score with regard to the constitutional law.<sup>12</sup>

### *Discussion*

The data collected in the second mock jury simulation study replicate the general pattern of results from the first simulation study, low accuracy overall for the jurors instructed with the MAIs, and lowest accuracy on the measure of procedural knowledge of state law. Across three independent studies, we have found that the approved instructions are not very effective at directing jurors to weigh aggravating and mitigating factors (Wiener et al., 1998, the current stud-

<sup>12</sup>Note that the mean accuracy differences between the pre-deliberation and post-deliberation measures for the pre- and post- and post-only groups are greater for procedural constitutional law than for declarative constitutional law, but the standard deviations (and therefore standard errors) are lower in the latter measures. This explains the seemingly differential *t* tests results reported in the text, that is, that the difference for declarative constitutional law is significant but it is not for procedural constitutional law.

ies). Further, we found that the weak deliberation effect that reached significance in the first simulation study (across all instruction groups) dropped out when we used only the MAIs and controlled for survey practice effects.

### Conclusion: Is There a Reasonable Likelihood of Confusion?

In *Boyde* (1990), the Supreme Court of the United States held that that there cannot be a “reasonable likelihood that the jury has applied the challenged instruction in a way that prevents the consideration of constitutionally relevant evidence” (p. 380). If there is a reasonable likelihood that the instructions prevented full consideration of all relevant information (such as the proper weight of mitigating and aggravating circumstances) a trial court that imposes the death penalty in a capital murder case denies the defendant his or her Eighth Amendment right to freedom from cruel and unusual punishment. If the jury does not understand how the law requires it to establish, weigh, and balance aggravation and mitigation, then it may well be requiring the defendant to forfeit his or her life without the benefit of due process of law. In short, the jury that acts with unbridled discretion is in danger of reaching a decision to impose the ultimate penalty without first considering the eligibility of the defendant or systematically considering whether the candidate meets the selection criteria for the punishment (*Buchanan*, 1998). In fact, recent research from a number of psychological laboratories suggests that the current jury instructions do not provide the kind of guidance that is necessary to reach a sentence that reflects both “rational orderliness” and “moral appropriateness” (Bilionis, 1993, p. 1653). Our own work, presented here in some detail, leads us to several conclusions about the jury instructions that jurors use in capital murder cases.

First, state-approved jury instructions may very well closely follow statutes and precedents set in the case law, but they are difficult for lay people, those who will serve as jurors, to understand. Mock jurors do not understand the totality of the penalty phase instructions (Wiener et al., 1995, 1998, and Jury Simulation Studies 1 and 2 in the current article). With regard to the Missouri instructions, eligible jurors do not understand the sentencing process any better after reading the instructions than they do after reading a control set of instructions that removes explanation of legal concepts and legal rules (Wiener et al., 1995; the current article). Second, jurors who do not understand the instructions (especially the procedures that make up the state penalty scheme) are more likely to assign a death sentence (even after a period of deliberation) than are jurors who do understand the instructions. We demonstrated this relationship in three experiments with three different samples (Wiener et al. 1995, 1998, and the current jury simulation results). Third, the lack of understanding that jurors bring to the sentencing task is, in part, a result of the common errors that they hold about homicide and the law. As we reported in the current article, jurors bring to the sentencing process errors in understanding some of the basic elements of law (i.e., the nature of aggravation and mitigation). Unless the instructions correct these errors, there is a great likelihood that jurors will make these same errors when they deliberate and assign a sentence. Our current data show that a debunking instruction designed to correct common errors may improve jury understanding of, at least, the concepts that make up the law.

Our current work tested successfully several methods of improving juror understanding of the concepts (declarative knowledge) and processes (procedural knowledge) in the law. These include the following: (a) Simplify the language of the instructions to improve comprehensibility of both the declarative and procedural information about the sentencing process; (b) present the instructions in a flowchart format so that jurors can understand the natural progression of law and the rules that the judge asks them to follow both as individuals and as juries; (c) allow the jurors an opportunity to review and practice the instructions in a mock jury case so that when it comes time to apply the law in a real case, they will be familiar with, at least, the general concepts and the flow of directions that make up the instructions; and finally, as mentioned above, (d) offer jurors debunking corrections to commonly held judgment errors.

Of course, some will argue that variations in the approved jury instructions will bias the jurors against one side or the other. The current data do not support such a conclusion (see also Wiener et al., 1995). In the regression analysis reported above in Jury Simulation Study 1, we tested the effects of jury attitudes, instruction type, and jury's accurate knowledge of the instructions in predicting sentencing outcome of 80 deliberating juries. Although we found a moderately strong relationship between accuracy of the juries' knowledge of the law and sentence, we found no difference in sentence outcome for the several instruction conditions. In other words, our fourth and final conclusion is that the influence that variations in instructions might have in murder cases most likely works indirectly by increasing understanding of the safeguards in the penalty phase process, and not by tilting the judgment process in one direction or the other.

Our research suggests that the penalty phase instructions in capital murder trials produce high levels of confusion in the minds of the jurors. Richard L. Wiener has made this argument in a number of cases in Missouri and Ohio. However, the courts have not yet been swayed by the data that others and we have presented to demonstrate juror confusion. The courts in Missouri (*Deck*, 1999) and in the Seventh Circuit (*Free v. Peters*, 1991) have rejected these data or other data like them because they found fault with the methodologies used to generate them. The current data and the data offered by others in this area use a diverse set of materials, samples, stimulus presentation, and simulated legal processes. Issues that have been put forth by the courts in rebuking the replicated results of studies such as the ones reported in this article center on the lack of control groups, the lack of deliberation, and the inability of most experimental materials to reflect the kinds of information that are presented at trial. We believe that the results of the current trial simulation experiments make these criticisms increasingly more difficult to put forth. Our current studies used video reenactments of actual trials in which professional actors played the roles of judges, attorneys, the defendant, and witnesses. Our stimulus trial included all the major components of a first-degree murder trial including opening and closing arguments by both attorneys, witness direct and cross examinations, and a judge instructing the jury. Death-qualified jurors deliberated and appeared to take the task very seriously.

Although no simulation can capture the experience of serving on a first-degree murder case, these simulations went a long way to speak to many of the issues raised in the *Free* (1991) and *Deck* (1999) cases. In fact, we collected a second set of simulation data to test specifically the effect of deliberation (a point made by

both the *Deck* Court and the *Free v. Peters* Courts). To our surprise, we found that deliberation had only a small positive effect on comprehension in our first simulation project and that when we removed the effects of practice and focused only on the approved jury instructions, we were unable to replicate the deliberation effect. In fact, in our second simulation, deliberation seemed to produce a slight detrimental effect on jury accuracy. Although it is true that the simulation process can never approximate an actual trial, the current generation of jury research projects is methodically removing the criticisms that the courts have levied in the past.

In response to the current work on jury compliance, one might argue that it is not possible to develop a penalty scheme that allows jurors to adequately consider aggravation and mitigation in a way that satisfies the demands for “rational orderliness” and “moral appropriateness” (Bilionis, 1993). If that were the case, then alternative sentencing procedures that bypassed jury sentencing (e.g., those that allowed judges to pass sentence on convicted capital murders) might be worth pursuing to ensure that capital offenders receive a careful hearing on the facts of their charges. However, the U.S. Supreme Court at least partially blocked this approach to solving the comprehension dilemma in its summer 2002 term, when it held in *Ring v. Arizona* (2002) that the Sixth Amendment guarantee of a trial by jury prohibits judges, alone, from finding aggravating circumstances necessary for imposition of the death penalty. The logic followed in *Ring* highlights the importance that the Court assigns to the way in which judges and jurors use aggravating and mitigating circumstances to reach penalty decisions.

In a prior case, *Walton v. Arizona* (1990) the Court held that because aggravating factors were “sentencing considerations” and not “elements of the offense” (p. 648), Arizona’s sentencing scheme did not curtail a defendant’s right to a jury verdict. The *Walton* Court held that Arizona’s statute asked judges to consider whether to “place a substantive limitation on sentencing,” the effect of which was to guide the “choice between life and death” and not to make findings of fact with regard to an “element of the crime of capital murder” (aggravation is not an element of first-degree murder in Arizona). However, in *Apprendi v. New Jersey* (2000), the Supreme Court ruled that a sentencing judge was prohibited from making a factual finding that the defendant’s crime “had been motivated by racial animus” (pp. 469–470), triggering New Jersey’s hate crime enhancement, which had the effect of doubling the defendant’s maximum prison sentence. In *Ring* (2002) the justices reasoned, “[I]f a State makes an increase in a defendant’s authorized punishment contingent on the finding of a fact, that fact—no matter how the State labels it—must be found by a jury beyond a reasonable doubt” (p. 2439). The Court said the issue here was “not one of form, but of effect” (p. 2430). Writing for the majority in *Ring*, Justice Ginsberg came to the conclusion that if the Sixth Amendment forbids a trial judge to decide an issue of fact that results in an enhancement of a penalty without a factual finding by a jury, then it follows that a judge sitting without a jury may not find aggravation in a capital murder case to impose the death penalty. Therefore, *Ring* overruled *Walton* “to the extent that it allows a sentencing judge, sitting without a jury, to find an aggravating circumstance necessary for imposition of the death penalty” (p. 2443).

The holding in *Ring* (2002) is a narrow one in that it makes illegal the sentencing schemes in the 5 states that assign the ultimate sentencing decision

entirely to judges (Arizona, Colorado, Idaho, Montana, and Nebraska; *Ring*, 2002, p. 2442) However, as Justice Scalia pointed out in his concurring opinion, the ruling in *Ring* seemingly does not preclude states from allowing judges to determine whether to impose the death penalty in cases where aggravation is made an element of the crime to be determined by the jury in the guilt phase of capital murder trials (*Ring*, p. 2445). Therefore, if a state required its juries to determine whether aggravation existed beyond a reasonable doubt in the guilt phase of capital murder trials, it could still leave the trial judges to weigh the aggravation against any mitigation to reach a final sentence and decide whether to impose the death penalty.

Recent research by Baldus, Woodworth, Young, and Christ (in press) suggests that judges in Nebraska weigh aggravating and mitigating factors in a balanced manner, which in fact may offset the finding that African American homicide defendants are more likely to face the possibility of the death penalty as a result of differences in prosecutors' use of their discretion in the urban, as opposed to rural, areas of Nebraska (i.e., urban prosecutors are more likely to seek the death penalty). Most African Americans in Nebraska live in urbanized counties (Baldus et al., in press). Therefore, in theory, it is possible to remedy juror inconsistency by increasing the role of judges in the states' capital murder sentencing process.

However, not all the justices in *Ring* (2002) embraced a narrow interpretation of the prohibition against judicial decision making in capital murder cases. Justice Breyer came to the same conclusion that the Constitution forbids judges to act alone to find aggravation that could trigger a death sentence by a very different route. He argued that "the Eighth Amendment requires that a jury, not a judge, make the decision to sentence a defendant to death." Justice Breyer concluded that its retributive function prevents the death penalty from constituting cruel and unusual punishment. He reached this conclusion by citing the empirical literature, which calls into serious doubt the effectiveness of the death penalty as a deterrent of homicide, by assuming that the death penalty is not necessary to incapacitate beyond the sentence of life in prison without the possibility of probation or parole, and by observing that rehabilitation is obviously not a purpose of the death penalty. Justice Breyer argued,

[With] respect to retribution, jurors possess an important comparative advantage over judges. In principle, they are more attuned to the 'community's moral sensibility . . . because they reflect more accurately the composition and experiences of the community as a whole.' (*Ring*, p. 2447)

Even if the U.S. Supreme Court does not endorse Breyer's minority opinion, it is likely that it will remain the accepted standard used in most of the jurisdictions in the United States.

Finally, in *Weeks* (2000), the U.S. Supreme Court looked to "empirical factors" to support the assumption that the jury follows the instructions it is provided (*Richardson*, 1987) and that it understands the answers that the judge provides to any of its questions (*Armstrong*, 1826). It is ironic that the empirical factors approach offered by Justice Rhenquist made no use of any of the many research studies cited in this article, given that the main business of psychology

and law scholars is to generate theory-driven, empirical arguments that speak to the legal issues and assumptions that the courts make (Wiener & Hurt, 1999). Had Justice Rhenquist taken seriously the psychological studies of juror comprehension in penalty phase instructions, he might have approached his analysis in the *Weeks* case somewhat differently. Rather than choosing to view instruction comprehension as an attribute of the jury, he might have considered both the characteristics of juries *and* the materials that the jury must comprehend in order to understand the law. It is important that empirical analyses take into consideration both the properties of the research participants, if by no other means than by controlling those properties through randomization, and the properties of the stimulus materials themselves. We believe that our work and the work of others demonstrate the properties of jury instructions that lead to low levels of jury comprehension. Our work and the work of others suggest ways in which the instructions can be improved to increase juror comprehension of the concepts and procedures that make up the charge in the penalty phase of capital murder trials. Clearly, more work needs to be done to understand how the instructions and jurors interact to produce varying levels of comprehension. In particular, psychologists have favored cognitive variables to the exclusion of motivational factors in explaining juror comprehension of instructions. Our future work in this area will begin to examine how motivational factors may facilitate and limit comprehension. We expect that new findings will tell us a great deal about the ways in which death penalty jurors “decide” to attend to some aspects of instructions and not others. In the meantime, we believe that our work and the work of others has extended a challenge to policy makers to improve the sentencing process so that the guided discretion of deliberating jurors grows from a legal fiction to an empirical reality.

### References

- Anderson, J. F. (2000). When the wall has fallen: Decades of failure in the supervision of capital juries. *Ohio Northern University Law Review*, 26, 2–40.
- Apprendi v. New Jersey, 530 U.S. 466 (2000).
- Armstrong v. Toler, 6 L. ED. 468 (1826).
- Ballew v. Georgia, 98 S. Ct. 1029 (1978).
- Baldus, D. C., Woodworth, G., Young, G. L., & Christ, A. M. (in press). The disposition of Nebraska capital and non-capital homicide cases (1973–1999): A legal and empirical analysis. *Nebraska Law Review*.
- Bilionis, L. D. (1993). Legitimizing death. *Michigan Law Review*, 91(7), 1643–1702.
- Blankenship, M., & Luginbuhl, J. (1994, March). *Juror comprehension of judicial instructions: A test of Tennessee’s death penalty process*. Paper presented at the meetings of the Academy of Criminal Justice Sciences, Chicago.
- Boyd v. California, 494 U.S. 370 (1990).
- Brown v. State, No. PCR2277 (St. Louis City, MO 1989).
- Buchanan v. Angelone, 522 U.S. 269 (1998).
- Butler v. State, No. 8610356SLC (Clay County, MO 1990).
- Cage v. Louisiana, 111 S. Ct. 328 (1990).
- Caldwell v. Mississippi, 472 U.S. 320 (1985).
- Charrow, R., & Charrow, V. (1979). Making legal language understandable: A psycholinguistic study of jury instructions. *Columbia Law Review*, 52, 386–407.
- Diamond, S. (1993). Instructing on death: Psychologists, juries, and judges. *American Psychologist*, 48, 423–434.

- Elwork, A., & Sales, B. (1985). Jury instructions. In S. Kassin & L. Wrightsman (Eds.), *The psychology of evidence and trial procedure* (pp. 280–297). Beverly Hills, CA: Sage.
- Elwork, A., Sales, B., & Alfini, J. (1982). *Making jury instructions understandable*. Charlottesville, VA: Michie.
- Free v. Peters, 778 F. Supp. 431 (N. D. Ill 1991).
- Furman v. Georgia, 408 U.S. 238 (1972).
- Gacy v. Wellborn, 994 F.2d 305 (7th Cir. 1993).
- Goodman-Delahunty, J., Greene, E., & Hsiao, W. (1998). Construing motive in videotaped killings: The role of jurors' attitudes toward the death penalty. *Law and Human Behavior, 22*, 257–271.
- Gregg v. Georgia, 428 U.S. 153 (1976).
- Haney, C., & Lynch, M. (1994) Comprehending life and death matters: A preliminary study of California's capital penalty instructions. *Law and Human Behavior, 18*, 411–436.
- Haney, C., & Lynch, M. (1997). Clarifying life and death matters: An analysis of instructional comprehension and penalty phase closing arguments. *Law and Human Behavior, 21*, 575–595.
- Haney, C., Sontag, L., & Costanzo, S. (1994). Deciding to take a life: Capital juries, sentencing instructions, and the jurisprudence of death. *Journal of Social Issues, 50*, 149–176.
- Hans, V. (1988). Death by jury. In K. C. Haas & J. A. Inciardi (Eds.), *Challenging capital punishment: Legal and social science approaches* (Sage criminal justice system annuals; Vol. 24, pp. 149–175). Thousand Oaks, CA: Sage.
- In re Winship, 397 U.S. 366 (1970).
- Kunda, Z. (1999). *Social cognition: Making sense of people*. Cambridge, MA: MIT Press.
- Lockett v. Ohio, 438 U.S. 586 (1978).
- Lockhart v. McCree, 106 S. Ct. 1758 (1986).
- Luginbuhl, J. (1992). Comprehension of judges' instructions in the penalty phase of a capital trial: Focus on mitigating circumstances. *Law and Human Behavior, 16*, 203–218.
- Lynch, M., & Haney, C. (2000). Discrimination and instructional comprehension: Guided discretion, racial bias, and the death penalty. *Law and Human Behavior, 24*, 337–358.
- McClesky v. Kemp, 481 U.S. 279, 308–09, (1987).
- McGautha v. California, 402 U.S. 183. (1971).
- Mills v. Maryland, 486 U.S. 367 (1988).
- Note. (2001). The rhetoric of difference and the legitimacy of capital punishment. *Harvard Law Review, 114*, 1559–1621.
- O'Neil, K., Patry, M., & Penrod, S. (2004). Exploring the effects of attitudes toward the death penalty on capital sentencing verdicts. *Psychology, Public Policy, and Law, 10*, 443–470.
- Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision-making. *Journal of Personality and Social Psychology, 51*, 242–258.
- Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 521–533.
- Richardson v. Marsch, 481 U.S. 200 (1987).
- Ring v. Arizona, 122 S. Ct. 2428 (2002).
- Severance, L. J., Greene, E., & Loftus, E. F. (1984). Toward criminal jury instructions that jurors can understand. *Journal of Criminal Law and Criminology, 75*, 198–233.
- Simmons v. South Carolina, 114 S. Ct. 2187 (1994).

- Smith, E. R. (1988). Category accessibility effects in a simulated exemplar-based memory. *Journal of Experimental Social Psychology*, 24, 448–463.
- Smith, E. R. (1989). Procedural efficiency: General and specific components and effects on social judgment. *Journal of Experimental Social Psychology*, 25, 500–523.
- Smith, E. R. (1990). Content and process specificity in the effects of prior experiences. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in social cognition* (Vol. 3, pp. 1–59). Hillsdale, NJ: Erlbaum.
- Smith, E. R. (1991). Illusory correlation in a simulated exemplar-based memory. *Journal of Experimental Social Psychology*, 27, 107–123.
- Smith, E. R., & Branscombe, N. R. (1987). Procedurally mediated social inferences: The case of category accessibility effects. *Journal of Experimental Social Psychology*, 23, 361–382.
- Smith, E. R., & Branscombe, N. R. (1988). Category accessibility as implicit memory. *Journal of Experimental Social Psychology*, 24, 490–504.
- Smith, E. R., Branscombe, N. R., & Bormann, C. (1988). Generality of the effects of practice on social judgment tasks. *Journal of Personality and Social Psychology*, 54, 385–395.
- Smith, E. R., & Lerner, M. (1986). Development of automatism of social judgments. *Journal of Personality and Social Psychology*, 50, 246–259.
- Smith, E. R., & Zarate, M. A. (1990). Exemplar and prototype use in social categorization. *Social Cognition*, 8, 243–262.
- Smith, V. L. (1991). Prototypes in the courtroom: Lay representations of legal concepts. *Journal of Personality and Social Psychology*, 61, 857–872.
- Smith, V. (1993). When prior knowledge and law collide: Helping jurors use the law. *Law and Human Behavior*, 17, 507–536.
- State v. Butler, 951 S. W.2d 600 (Mo. 1997).
- State v. Deck, 994 S. W.2d 527 (Mo. 1999).
- State v. Jones, 979 S. W.2d 171 (Mo. 1998).
- State v. Knese, 985 S. W.2d 759 (Mo. 1999).
- Strawn, D. J., & Buchanan, R. W. (1976). Jury confusion: A threat to justice. *Judicature*, 59, 478–483.
- Walton v. Arizona, 497 U.S. 639 (1990).
- Wainwright v. Witt, 469 U.S. 412 (1985).
- Weeks v. Angelone, 120 S. Ct. 727 (2000).
- Wiener, R. L. (1993). Social analytic jurisprudence and tort law: Social cognition goes to court. *Saint Louis University Law Journal*, 37, 503–551.
- Wiener, R. L., Habert, K., Shkodriani, G., & Staebler, C. (1991). The social psychology of jury nullification: Predicting when jurors disobey the law. *Journal of Applied Social Psychology*, 21, 1379–1401.
- Wiener, R. L., & Hurt, L. E. (1999). An interdisciplinary approach to understanding social sexual conduct at work. *Psychology, Public Policy, and Law*, 5, 556–595.
- Wiener, R. L., Hurt, L. E., Thomas, S. L., Sadler, M. S., Bauer, C. A., & Sargent, T. M. (1998). The role of declarative and procedural knowledge in capital murder sentencing. *Journal of Applied Social Psychology*, 28, 124–144.
- Wiener, R. L., Pritchard, C. C., & Weston, M. (1995). Comprehensibility of approved jury instructions in capital murder cases. *Journal of Applied Psychology*, 80, 455–467.
- Wiener, R. L., Richmond, T. L., Seib, H. M., Rauch, S. M., & Hackney, A. A. (2002). The psychology of telling murder stories: Do we think in scripts, exemplars, or prototypes? *Behavioral Sciences and the Law*, 20, 119–139.
- Wiener, R. L., Wiener, A. T., & Grisso, T. (1989). Empathy and biased assimilation of testimonies in cases of alleged rape. *Law and Human Behavior*, 13, 343–355.
- Witherspoon v. Illinois, 391 U.S. 510 (1968).