Psychology, Crime & Law (1996) Vol. 2, pp 269–297
Reprints available directly from the publisher
Photocopying permitted by license only

FACTORS AFFECTING SIMULATED JURORS' DECISIONS IN CAPITAL CASES[1]

VLADIMIR J. KONEČNI, EBBE B. EBBESEN, and ROGER R. HOCK

University of California, San Diego

Previous psycho-legal research has claimed that the process of selecting death-qualified jurors for capital cases creates conviction-prone juries. The studies on which these claims are based have employed simulation methodologies to examine the relationship between subjects' death-penalty attitudes and verdict decisions, as well as the effect of the death-qualifying *voir dire* itself. Despite admitted weaknesses of simulations in general, this method was employed in the present research so that conceptual comparisons to past findings could be drawn. Two experiments were designed to examine the issue of death-qualification and biased juries in a context of other potentially highly influential factors, namely, the strength of evidence and the degree of heinousness. Our results failed to find any of the relationships between death-penalty attitudes and verdict decisions that would be predicted from past research. Instead, the subjects' decisions were influenced, virtually exclusively, by the strength of the evidence presented in the case, as is legally prescribed. In the light of these findings, the discussion focused on the questions of reliability and external validity of simulation research, the potential problems caused by method-specific factors in determining the outcome of such methodology, the attitude-behavior link, and the danger of premature and unwarranted application to the legal system of findings from simulations.

Key words: capital punishment; "death-qualified" juries..

In the legal system of the United States, veniremen in capital cases can be excluded from serving on a jury if a judge decides that they hold strong beliefs opposed to the death penalty. The rationale for this exclusionary rule has been that these citizens, if empaneled, may prevent the legal imposition of the death penalty either by refusing to find a defendant guilty or by refusing to vote for capital punishment in the sentencing phase of the trial, regardless of the merits of the case. All cases, therefore, in which the possibility of capital punishment existed, have been decided by "death-qualified" juries. In *Witherspoon v. Illinois* (1968), the death penalty imposed on the defendant was challenged on the grounds that the systematic exclusion of jurors opposed to the death penalty had created a jury biased toward bringing in the verdict of "guilty." It was argued that the exclusive use of death-qualified jurors violated Witherspoon's constitutional rights to a fair trial as guaranteed by the 6th and 14th Amendments.

The few, then unpublished, social-science studies presented to the U.S. Supreme Court when it heard the case (Wilson, 1964; Zeisel, 1968; Goldberg, 1970) failed to convince the Court that Witherspoon had been convicted by a biased jury. In a split decision, the Court found the evidence supporting this contention to be of a "tentative and fragmentary" nature and that Witherspoon had not met the burden of proof that jurors who are not opposed to capital punishment are more prone to convict. This precedent-setting decision confirmed the constitutionality of the use of death-qualified juries and opened the door for subsequent research on the issue of systematic bias of juries thus empaneled.[2]

All veniremen are questioned about their attitudes and feelings regarding case-related issues prior to being empaneled (*voir dire*). Those who in the opinion of the judge would for

any reason be unable to render an impartial verdict are excused from serving. Juries in capital cases are unique in that they have the potential dual responsibility of deciding both the guilt of the defendant and the sentence. Figure 1 is an authentic representation of the chart used in the Superior Courts of San Diego to explain to prospective jurors in capital cases the extent of their decision-making responsibilities should they be empaneled. Therefore, the *voir dire* of such prospective jurors includes the additional process of "death qualification."

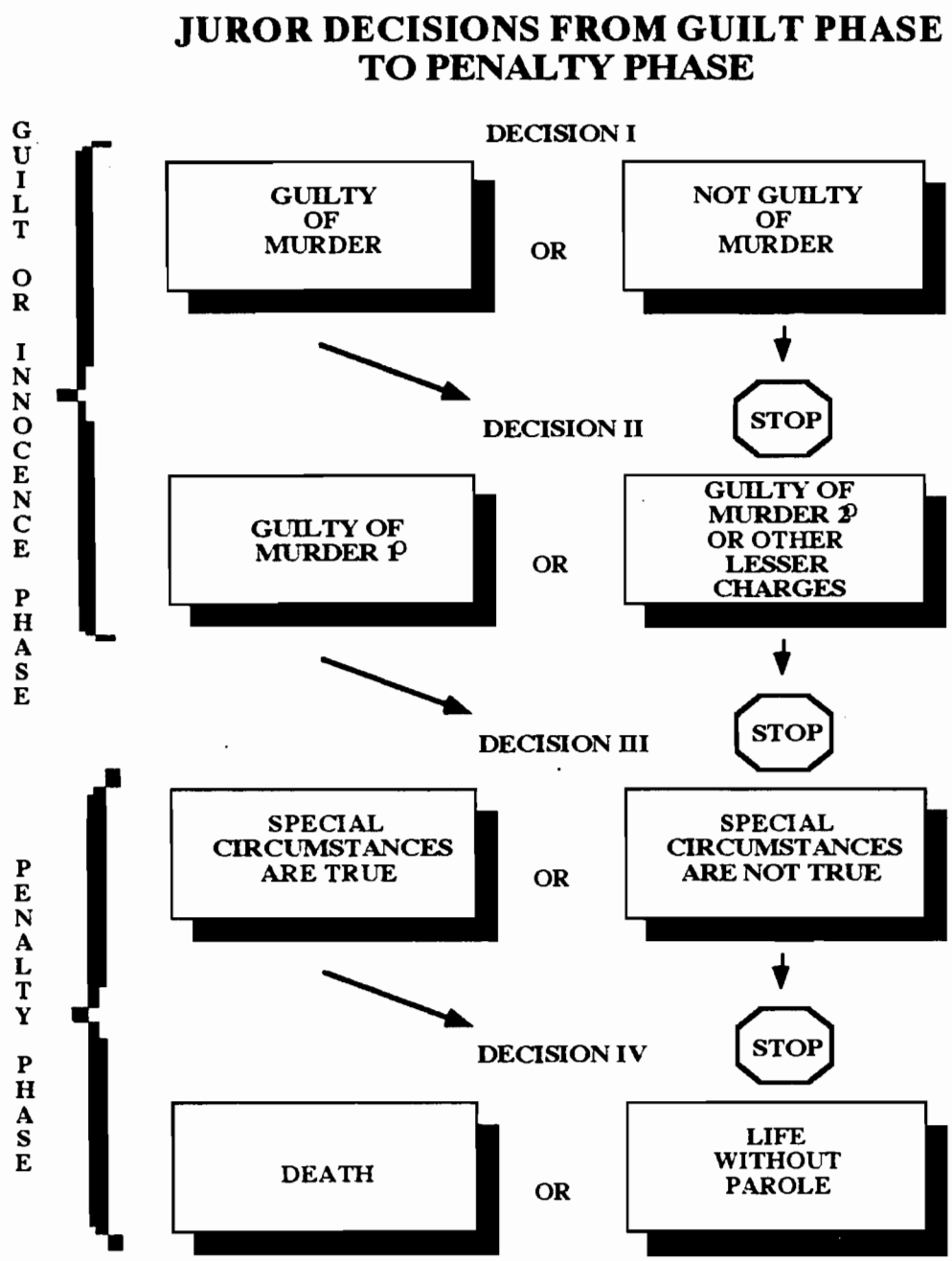
In capital cases following *Witherspoon*, the goal of the death-qualification portion of *voir dire* is to identify three groups of veniremen: *Nullifiers*, who would be unable to render an impartial verdict due to their beliefs about the death penalty and are therefore excused from serving; the so-called *Witherspoon* excludables, who would refuse to vote for the death penalty regardless of the evidence presented (even though they could render an impartial verdict), and are thus also excused; and the so-called *death-qualified* veniremen who fall into neither of the above categories and are deemed qualified to sit on the jury in a capital case.

It is the systematic exclusion of veniremen who oppose the legally available death-penalty option regardless of the merits of the case, their verdict-reaching impartiality notwithstanding, that has given rise to allegations that the composition of juries in capital cases is skewed toward conviction-proneness. Since "nullifiers" would be excluded from serving regardless of the *Witherspoon* decision, the key research comparison has been between *Witherspoon* excludables and death-qualified jurors.

Early research on the issue of death-qualification and jury bias focused on the relationship between death-penalty attitudes and other attitudes toward crime and punishment (Bronson, 1970; Goldberg, 1970; Zeisel, 1970; Jurow, 1971). These studies have been reviewed elsewhere (e.g., Haney, 1984; Thompson, Cowan, Ellsworth and Harrington, 1984) and need not be discussed again here. The study by Moran and Comfort (1986) is the only one to date which found a positive relationship between pro-capital-punishment views and conviction-proneness in people who had previously, in fact, been impaneled on a felony jury. However, not only was a very low percentage of mailed questionnaires returned (22.7%), but the failure to provide a breakdown by type of felony raises the distinct possibility that the results are irrelevant for verdicts in capital cases.

Other research has focused on the specific question of whether death-qualified simulated jurors as a group are biased toward conviction in comparison to the group of potential jurors who are excluded under the *Witherspoon* decision. Cowan, Thompson and Ellsworth (1984) divided subjects into nullifiers, *Witherspoon* excludables, and those who were considered death-qualified. This was done by asking one question about the subjects' ability to render an impartial verdict, given their attitudes toward the death penalty (to identify the nullifiers), and one question concerning their willingness to impose the death penalty (to identify the *Witherspoon* excludables). After eliminating the nullifiers from the sample, the remaining 30 excludable and 258 death qualified subjects were shown a videotaped reenactment of a synopsized murder trial. On several measures, the death-qualified group was significantly more likely to convict than the *Witherspoon* excludables.

Thompson et al. (1984) presented death-qualified and *Witherspoon* excludable subjects with a videotaped simulation of an assault so as to assess the possible differences in the way the two groups interpret the evidence presented. Death-qualified subjects interpreted testimonial evidence as more favorable to the prosecution than did the excludable group. The



Used by permission: Office of the District Attorney, San Diego, California

Figure 1. San Diego Superior Court chart illustrating jury decision process in capital cases.

authors concluded that the alleged difference between the two groups in conviction-proneness may be due, in part, to differential perceptions of the strength of the evidence against the defendant.

A related line of research examined the possibility that the death-qualifying *voir dire* procedure itself could have a biasing effect on the subsequent decisions of jurors. Haney (1984) asked a group of subjects (from which nullifiers and excludables had been removed) to watch a videotape of a *voir dire* procedure in a capital case being performed on other simulated jurors. For half of these observer-subjects, the tape included the death-qualifying process, while for the other half this part of the tape was edited out. All subjects then completed a questionnaire. Findings indicated that subjects who had watched the death-qualifying questions and answers were more likely to (a) believe that the defendant would be found guilty, (b) believe that the judge and both attorneys assumed the defendant to be guilty, and (c) choose the death penalty than the subjects who had watched the jury selection without the death-qualifying segment.

Although no case of any kind was presented and no verdicts were rendered, this single study (at that time unpublished) was cited as the key factor in a California Supreme Court decision to require individual sequestration of prospective jurors for the death-qualifying portion of *voir dire*. The opinion written by then Chief Justice Rose Bird,[3] stated:

> "In order to minimize the potentially prejudicial effects identified by the Haney study, this court declares, pursuant to its supervisory authority over California criminal procedure, that in future capital cases that portion of the voir dire of each prospective juror which deals with issues which involve death-qualifying the jury should be done individually and in sequestration" (*Hovey v. California*, 1980).

This opinion demonstrates the potential of social-science research to influence in important ways the fundamental workings of the criminal justice system. The *Hovey* decision changed the way capital juries are selected in California, adding, at minimum, cost and time loss to the process. While this kind of direct impact has been rare, a considerable problem is demonstrated here in applying findings from simulation research, the external validity of which had yet to be established. Although the various studies described here support the notion that death-penalty attitudes may affect verdict decisions, and therefore allow the inference that conviction-biased juries are being created through the use of *Witherspoon* exclusions, we find this argument far from convincing, given the nature of the research that has been conducted in this area. [See the complementary criticism by Elliott (1991a) of the brief the American Psychological Association filed in *Lockhart v. McCree* (1986).]

Most of the above research (and all the recent studies) has relied on simulation methodologies. The difficulty of achieving external validity with the exclusive use of this method is well-known (see Konečni and Ebbesen, 1979, 1982). There is little reason for assuming *a priori* that survey questions, written case-summaries, audio- or video-tapes can produce attitudes in subjects that activate or reflect the same underlying decision mechanisms that control the behavior of real juries.

Furthermore, these studies are inferentially weak for an even stronger reason, namely, their exclusion of other presumably important influences on real jurors' verdicts. The criminal justice system predicates that verdicts be based solely upon the facts of the case as presented during the trial—the evidence. Even if prior research has consistently shown the effects of death-penalty views on verdicts, it should have been self-evident to those research-

ers that the relative importance of their effects could not be known until the percent of variance these effects accounted for was compared to that accounted for by the evidence. If inferences about the legal system are to be drawn from simulation research, it is crucial to simulate other important factors in the real-world system that can be plausibly believed to influence the behavior of interest. The influence of pretrial attitudes may, for example, be eliminated when the evidence and the jurors' motivation to follow the judge's instructions to base their verdict solely on that evidence are both strong.

We therefore carried out two simulation studies that examined the relationship of death-penalty attitudes to verdicts in the context of an experimental manipulation of the strength of two key trial factors: evidence and heinousness. Although two of the present authors have written extensively on the potential dangers of generalizing from simulation methodology in psycho-legal research (e.g., Ebbesen and Konečni, 1980, 1985; Konečni and Ebbesen, 1979, 1981, 1982, 1984, 1992), we intentionally chose this methodology in the present research to allow for comparisons (conceptually, at least) between our findings and the existing body of literature. It might seem strange that a study designed to explore the external validity of previous simulations would continue to use simulation methodology. One might rightly argue that a better test of generalizability would employ methods very different from those already used. However, in the present case, because we were on various grounds predicting results different from those previously found, we felt that a stronger argument could be made about the inferential errors made by the authors of prior studies if our procedures were methodologically similar to those they utilized. Were our methods completely different from those used in past research, critics could point to these differences as explanations for the failure to replicate.

The following factors were combined in a single experimental design.

Evidence. The evidence presented in open court during a trial ought to be the primary and central factor in a jury verdict. In order to examine this factor in comparison with others that are functioning concurrently in a legal setting, the strength of the case against the defendant was varied systematically.

Heinousness. This is a factor present in all capital cases, but its effect on simulated jurors' verdicts has not before been studied. It was included in order to determine what influence, if any, it may exert on verdicts in the context of a simulation-research approach. Heinousness was also used to provide a test of discriminant validity in that, while heinousness might not affect verdicts, it should (if our procedures and methods are adequately sensitive) influence sentencing decisions.

Voir dire. Claims have been made (Haney, 1984) and accepted by the California Supreme Court in *Hovey* that the very process of questioning veniremen about the death penalty prior to being empaneled may bias them toward guilty verdicts. However, due to serious questions about the adequacy and completeness of the prior research, this contention was re-examined using a different approach than had been employed in previous work.

Death-penalty attitudes. The connection between death penalty attitudes and verdicts in capital cases is the central issue in the death-qualification controversy. Until now, however, this connection has not been tested in the presence of variations of the other key factors listed

above. It can be argued from reasoning about stimulus-person interactions (Mischel, 1976), for example, that death penalty attitudes will only exert their effects in cases in which the evidence against the defendant is weak.

Since evidence was found to be the overwhelmingly dominant factor in the subjects' verdicts in the first study, a second experiment was designed to replicate Experiment 1 conceptually and to explore the influence of evidence in greater detail.

EXPERIMENT 1

METHOD

Subjects, Design and Materials

Undergraduate students (N = 198) at the University of California, San Diego who indicated that they either had a valid California driver's licence or were registered to vote in California were randomly assigned to groups exposed to simulated murder cases involving varying amounts of evidence (three levels), differing degrees of heinousness (two levels), and differing death-qualifying *voir dire* conditions (two levels), in a between-subjects design. The subjects' assessed levels of scruples regarding the death penalty was treated as a co-variate.

Six written case summaries of capital crimes were developed. They varied systematically in terms of evidence and heinousness. All case summaries were variations of the same basic crime: A murder committed by a handgun during the commission of a robbery in a convenience store. Both the prosecutor's and the defense attorney's cases were summarized.

Evidence. In the weak evidence condition, there was one witness who was hiding in the store and heard, but did not see, what took place. She was able to make a voice identification and offer testimony as to the defendant's attire. Other relatively weak circumstantial evidence was presented. In the moderate evidence condition, the witness hiding in the store had a better view of the robbery and shooting and was able to make a positive identification of the defendant and offer more detailed information about his attire. Additional evidence regarding cash found in the defendant's possession was also included. In the strong evidence condition, there were two eyewitnesses hiding in the store. Their accounts of the events of the murder agreed, including their identification of the defendant and his attire on the night of the crime. Additionally, in the strong evidence condition, the murder weapon, wrapped in a jacket possibly belonging to the defendant, was found in a trash dumpster near the place where the defendant was captured. A pilot study determined that these three levels of evidence were at approximately equal intervals of discriminability.

Heinousness. In the low heinousness condition, the gunman entered the store and demanded money. The store attendant, while complying, triggered the store alarm at which time the robber shot the attendant, grabbed the money and ran out of the store. In the high heinousness condition, the gunman entered the store, pistol-whipped the attendant repeatedly, tied him up, forced him to kneel, and shot him. He then shot the body of the attendant five more times before grabbing the money and running out of the store. A pilot study determined that subjects clearly discriminated between these levels of heinousness.

The various levels of evidence and heinousness were combined to create six different case summaries (see Appendix 1 for a sample case summary).

Voir Dire and Death-Penalty Scruples. A 10-item questionnaire was developed to assess the subjects' scruples about the death penalty. In an attempt to insure that this simulated *voir dire* asked legally appropriate questions about death-penalty beliefs, the first five items were stated exactly as they are in the preliminary screening of jurors in the San Diego Superior Court system in cases where the death penalty is being sought by the prosecution. The final five items were designed to be similar to some of the follow-up questions which the attorneys or the judge might ask prospective jurors during *voir dire* in order to assess further their position on the death penalty (see Appendix 2 for the complete questionnaire). This questionnaire was given to all subjects in the study. However, half of the subjects responded to the questionnaire prior to reading the case summary (pre-*voir dire* condition), while the other half answered the questionnaire after they had been exposed to the case summary and reached the decisions about the verdict and sentence (post-*voir dire* condition).

This death penalty questionnaire was also used to measure the subjects' scruples regarding the death penalty (the fourth factor in the design). Thus a single instrument was employed both to assess the death-penalty scruples and, by varying when the subjects completed the questionnaire, to examine the effect of pre-trial exposure to a simulated death-qualifying procedure.

Procedure

Subjects were treated in groups of 12 to 15. Each subject was randomly assigned to one of 12 conditions combining evidence, heinousness, and pre- versus post-*voir dire*. Each subject received a packet which contained the following: 1) case summary; 2) verdict measure; 3) special circumstances measure; 4) sentence measure; 5) death penalty questionnaire; and 6) demographic items. As mentioned above, the death penalty questionnaire came first in the packet for half of the subjects and after the sentence measure for the other half.

After being instructed about their role as jurors, subjects were given simulated California Jury Instructions Criminal (CALJIC) instructions that emphasized the need to base their decisions solely on the evidence. Subjects were then told to proceed through their packets slowly and carefully, one page at a time, and not to look ahead in the packet at any time. Once they had responded to an item, they were not allowed to change their answers.

The verdict measure required subjects to find the defendant guilty or not guilty of first degree murder and to indicate the level of confidence in the verdict decision. The special circumstances page explained the meaning of this concept and required the subject to decide if these circumstances were true for the case presented. For the sentence measure, subjects were required to decide between the death penalty and life imprisonment without the possibility of parole. Only those who found the defendant guilty were required to make a decision regarding special circumstances and sentence. Demographic items were: Gender, age, whether the subject had a California driver's license, whether the subject was registered to vote in California, and the subject's political party affiliation. Voter registration and driver's license lists are commonly used for selecting individuals for jury duty in California.

RESULTS

Scores on the death-penalty attitude measure were coded so that higher scores indicated increasingly pro-capital punishment attitudes. One of the original ten items was dropped from the total since it was designed to assess impartiality and therefore did not directly relate to the attitude score. The death-penalty measure was found to have a fair degree of internal consistency when tested overall (Alpha = .73). On the other hand, when the five items taken from the San Diego Superior Court were analyzed independently from the five new follow-up items, the Court's items were found to contain poor internal consistency (Alpha = .04), probably due to the confusing wording of the questions. The new follow-up items fared considerably better (Alpha = .76).

To enhance the external validity of the scruples measure, responses to all nine questions were used to compute total scruples scores, despite the lack of internal consistency in the five questions used by the San Diego Court. These total scores on the questionnaire were divided into equal thirds to produce three groups of subjects—holding pro-, moderate, or anti-death-penalty attitudes.[4] This trichotomized-scruples variable was used in some of the subsequent data analyses.

Table 1 Chi-Square and Log-Linear Analyses of Evidence (E), Heinousness (H), Pre- vs. Post-Voir Dire (P/P), and Death-Penalty Scruples (DP) in Relation to Sentence (S)

Factor	*df*	χ^2	*p*
	χ^2 Main Effect Analyses on Sentence		
Evidence	2	1.80	> .10
Heinousness	1	4.12	.05
Pre/Post Voir Dire	1	2.80	< .10
Death Penalty Scruples	2	24.03	< .001
	Model[1]		
Marginals + (S × DP)[2]	21	17.87	.66
Marginals + (S × DP) + (S × E)	19	16.07	.65
Residual from dropping (S × E)	2	1.87	> .10
Marginals + (S × DP) + (S × H)	20	13.68	.85
Residual from dropping (S × H)	1	4.19	< .05
	Model of Best Fit		
Marginals + (S × DP) + (S × H)	20	13.68	.85

[1]Model analyses were collapsed across Pre/Post Voir Dire for log-linear analysis to avoid excessive zero-cells. Results were unchanged when analyses were performed on a data table collapsed across evidence and the contribution of Pre/Post Voir Dire was examined.

[2]The (S × Factor) parameters in the log-linear models are the equivalent of main effects in analysis of variance models in which sentence is the dependant variable.

In order to assess the efficacy of the death-penalty scruples measure, analyses were performed treating sentence as the dependent variable, and using the trichotomized scruples variable, heinousness, pre/post *voir dire* and evidence as predictors. To the extent that our scruples-measurement and case-simulation procedures tapped attitudes and decision processes similar to those active in real jury decisions, it was logical to expect scruples and heinousness to predict sentence as they might in the penalty phase of a real trial. Moreover, unless such relationships were found, it could be argued that any failure of scruples to predict *verdict* could have been due to an invalid or poorly designed death-penalty attitude measure. As can be seen in Table 1, the connection between scruples and sentence was as expected. Chi-square analyses of subjects who found the defendant guilty indicated that death-penalty scruples and heinousness were the only factors, when considered independently, that were significantly related to sentence: Death sentences were more likely the more heinous the crime, and the more pro-death the subjects' attitudes. (In fact, not one of the subjects classified as anti-death chose the death penalty, but 66.7% of those classified as pro-death selected the death penalty.) This is the same configuration of factors that one would expect to have a combined effect on sentencing decisions in real-life capital trials. Furthermore, the fact that the scruples measure was a far better predictor of sentences than any of the case variables strongly suggests that the scruples measure is a valid one.

The extent to which these and other factors were able to significantly improve our ability to predict sentence choices was assessed by adding parameters estimating the interaction of each factor with sentence in various log-linear models. These log-linear odds analyses indicated that only the scruples × sentence and heinousness × sentence parameters were necessary to fit a model to a four-way data table from which the pre-post factor was dropped. In addition, none of the two-way interactions among the predictors improved the fit of the "main-effects-only" model.[5]

Such strong evidence that the scruples measure was validly assessing death-penalty attitudes permitted us to analyze the effects of these same factors on verdict decisions. Table 2 summarizes the analyses carried out to assess the influence of evidence, heinousness, pre- versus post-trial *voir dire* administration, and death-penalty scruples on guilty/not-guilty verdict decisions. Chi-square analyses were followed by a log-linear odds analysis to assess the relative fits of different models to the data.

Taken together, the results clearly suggest that the only factor meaningfully related to verdict was evidence. The extent to which other factors affected verdict decisions was assessed by examining whether the addition of parameters estimating the interaction of each factor with verdict significantly improved the fit of log-linear models to a four-way data table. Although the heinousness factor was dropped from the log-linear analysis to avoid excessive empty cells, a similar analysis in which the pre/post *voir dire* variable was dropped and heinousness was included yielded virtually identical results. It is clear from Table 2 that only the evidence × verdict parameter was required to fit the data. The addition of death-penalty scruples, or pre- vs. post-trial *voir dire* (or heinousness) failed to improve significantly the model's fit.

In what some might argue is a more direct test of the relationship between *Witherspoon* excludability and verdicts, death-qualified subjects were separated from *Witherspoon* excludables (and nullifiers were dropped) using appropriate items from the questionnaire. This

Table 2 Chi-Square and Log-Linear Analyses of Evidence (E), Heinousness (H), Pre- vs. Post-Voir Dire (P/P), and Death-Penalty Scruples (DP) in Relation to Verdict (V).

Factor	*df*	χ^2	*p*
	χ^2 Main Effect Analyses on Verdict		
Evidence	2	33.22	< .001
Heinousness	1	.56	.46
Pre/Post Voir Dire	1	.84	.36
Death Penalty Scruples	2	1.91	.38
	Log-Linear Models		
Marginals + (V × E)	27	20.93	.79
Marginals + (V × E) + (V × DP)	25	19.04	.80
Residual from dropping (V × DP)	2	1.89	> .50
Marginals + (V × E) + (P/P × DP)	25	18.23	.83
Residual from dropping (P/P × DP)	2	2.70	> .30
	Model of Best Fit		
Marginals + V × E	27	20.93	.79

[1]Fits to log-linear models were collapsed across heinousness to avoid excessive zero-cells.

variable (juror status) was substituted for the trichotomized-scruples factor and examined using a Chi-square analysis in relation to the verdict decisions. No significant relationship was found ($\chi^2(1) = .09$, $p > .90$; heinousness was dropped from the log-linear analysis to avoid an excess of empty cells, but a similar analysis in which heinousness was included and the *voir dire* factor dropped yielded virtually identical results). Table 3 contains the verdict choices of all simulated jurors, excluding the nullifiers, as a function of strength of evidence and juror status. As can be seen, over all cases, simulated *Witherspoon* excludables were no more likely to convict than were includables and somewhat counter to our initial expectations, the effect of juror status on the odds of a guilty verdict was unaffected by the strength of

Table 3 Verdict Decisions for Simulated Jurors as a Function of Strength of Evidence and Juror Status (Death Qualified vs. Excluded)

	Strength of Evidence					
	Weak		*Moderate*		*Strong*	
	Juror Status					
Verdict Decision	*Qualified*	*Excluded*	*Qualified*	*Excluded*	*Qualified*	*Excluded*
Guilty	9	3	24	10	31	4
Not Guilty	23	8	12	3	7	2

the evidence. Finally, also counter to initial expectations, a similar lack of effect was found for different levels of heinousness.

Additional analyses examined the relationships between the subjects' gender and political party affiliation and their scruples about the death penalty. Chi-square analyses demonstrated a marginal relationship for gender ($\chi^2(2) = 4.44$, p. = .10) and for political party affiliation ($\chi^2(8) = 14.65$, $p = .06$) with males and Republicans being slightly more likely to express pro-death attitudes than females and Democrats. However, when these same demographic variables were analyzed in relation to verdict, no significant relationships were found.

DISCUSSION

The criminal justice system rests largely on the assumption that the central factor influencing jurors' verdict decisions is the evidence presented in open court. Findings from past experiments concerning the verdicts of death-qualified juries in capital cases have suggested that the simulated jurors' attitudes about the death penalty also may be influencing their verdict decisions in important ways. If such findings are accurate and generalizable to the real world of courts, then the present system of jury selection for capital cases ought indeed to be challenged.

The results from the present study, however, strongly contradict these past findings. If we are to assume—an assumption that has been readily granted to previous simulations by both social scientists and the California Supreme Court—that our findings are accurate and generalizable (a claim that we, however, hesitate to make), our results could be taken as substantial support for the present process of jury selection with regard to the exclusion of the *Witherspoon*-defined group. Additionally, because there was no pre- vs. post-*voir dire* effect, the present findings completely failed to support the mandate created by the California Supreme Court in *Hovey* to sequester jurors during death-qualifying *voir dire*.

Our results indicated that evidence was the only factor that meaningfully influenced the subjects' verdict decision. Whether subjects strongly favored, strongly opposed, or held moderate beliefs regarding the death penalty did not significantly affect their verdicts. Not surprisingly, and lending additional credence to our instruments and measurements, the death-penalty attitudes strongly influenced whether a subject would impose the death penalty after a guilty decision had been made. In addition, whereas the heinousness of the crime was not related to the verdict decisions, just as is legally appropriate, it did influence the sentencing decisions, again, as is specified by law.

Furthermore, when those subjects who would be considered excludable under *Witherspoon* were separated and compared to those identified as death-qualified, the verdicts of the two groups were virtually identical.

Exposure to the death-qualifying *voir dire* questions before deciding a case had no effect on the subjects' verdicts. It is notable and instructive that while exposure to a simulation of the death-qualification process may influence a person's attitudes toward a case, as reported by Haney (1984), the actual decision-making *behavior* of simulated jurors is determined by factors other than the death-qualifying procedure, such as evidence.

We do not claim to have replicated Haney's study, nor to have made an attempt to do so. Haney's work focused on the effects of *observing* the death-qualification *voir dire* undergone by others and the possibility of being somehow changed by that observation (a "vicarious" process effect). Our research was concerned with what may be called a "direct" process effect, where the subjects are *participants* in *voir dire* and must make individual decisions about their own attitudes regarding the death penalty. It could and should be argued that this direct confrontation with the issue of the death penalty ought to have had even stronger effects than the observation of others being questioned (cf. Tesser, 1978). Yet, when rendering verdicts, our subjects were not affected by their participation in a simulated death-qualification *voir dire*. Issues of comparative realism aside, the point here is that there are many potential "process effects" of death-qualification that may or may not ultimately affect the verdict decisions. The only way to test for these effects is to provide subjects with the opportunity to render a verdict *after hearing the evidence*. Unfortunately, in Haney's (1984) research that had such an impact on judicial decisions subjects were never presented with trial evidence, nor were they asked to reach verdict decisions.

Researchers (e.g., Cowan et al., 1984) have argued that the systematic exclusion of anti-death penalty individuals from capital juries may produce "unrepresentative" juries because certain groups within the population (i.e., women and minorities) are disproportionately excluded. Our findings offer weak yet noteworthy support for the contention that women were overrepresented in the excludable group. Significantly, though, we found no difference between males and females (in our admittedly limited student population) in the *verdicts* they rendered.

EXPERIMENT 2

In order to obtain further comparisons with prior research and because we suspected that the extent to which scruples would predict verdicts might depend on the *nature* of the evidence in the case, a second experiment was designed. A somewhat different scruples measure was employed and wider variations in the nature of the evidence in the case were constructed to increase the range of stimulus conditions used.

METHOD

Subjects, Design and Materials

Subjects for the second study were 179 undergraduate students at the University of California, San Diego who had valid California driver's licences or were registered to vote in California. Participation in research projects was part of their course requirements.

Summaries of murder cases were written exactly as in Experiment 1 except that evidence was factored into three levels of eyewitness evidence and three levels of physical evidence (weak, moderate, and strong) on the assumption that a scruples × evidence interaction may depend on the nature as well as strength of the evidence. The eyewitness conditions involved increasing visibility and number of eyewitnesses as in Experiment 1. Strength of physical evidence was varied by increasing the quantity and decreasing the circumstantial nature of

the evidence (see Appendix 1 for a sample case summary). This created a 3×3 between-subjects design. In view of the results of Experiment 1, heinousness was not varied in Experiment 2.

Since the five items on the death-penalty questionnaire that were taken from the San Diego Superior Court were internally inconsistent and, we concluded, difficult to understand, these were eliminated from the questionnaire in this second study. Instead, five new questions were developed and combined with the five "follow-up" questions from our first study to create the second questionnaire (see Appendix 2 for the complete questionnaire).

Procedure

Subjects were again tested in groups of 12 to 15. Each subject was randomly assigned to one of the nine conditions and presented with a packet containing the death penalty questionnaire, the case summary, a verdict measure, and several demographic items as in Experiment 1. Since the murder described in the case summaries was committed during a serious felony (armed robbery), the presence of special circumstances was assumed and the subjects were not asked to decide this issue. Another reason for this was that nearly all the subjects in Experiment 1 who found the defendant guilty also found the special circumstances to be true due to the nature of the crime.

RESULTS

As expected, the death-penalty attitude questionnaire developed for this second study was found to have a higher degree of internal item-consistency (Alpha = .85) than the original questionnaire. Scores on the questionnaire were trichotomized as in Experiment 1 and analyzed along with eyewitness and physical evidence in relation to sentence and verdict. In a manner identical to the first experiment, the revised measure of scruples was significantly related to sentence ($\chi^2(2) = 33.10$, $p < .001$). As would be expected if the scruples measure were valid, only one of the 33 subjects classified as having anti-death attitudes chose the death penalty, while 61% of the 28 pro-death subjects chose this sentence. The revised scruples measure thus received strong support. Also analogously to Experiment 1, neither of the two evidence factors (nor their interaction) had a significant effect on sentence.

Whereas the scruples measure predicted the sentences the subjects imposed, only the eyewitness-evidence and the physical-evidence factors predicted the verdicts, as can be seen in Table 4. Both evidence factors were necessary in a log-linear odds analysis to describe the data. However, the scruples variable was again unnecessary and did not improve the fit of the model. The interaction between the two types of evidence did not approach significance. Table 5 reflects the arrangement of the data for the model of best fit. Replicating the results of Experiment 1, additional analyses found no significant interactive effect on verdicts between scruples and either of the evidence factors. As in the first study, subjects were again divided into death-qualified and *Witherspoon*-excludable groups according to the pattern of their answers to the questionnaire and compared on verdicts. And again no differences in verdict choices were found between these two groups ($\chi^2(1) = .089$, $p = .96$).

Experiment 2 successfully replicated the subjects' utter reliance on the strength of evidence (both eyewitness and physical) in reaching their verdicts. Also, the lack of a relation-

Table 4 Chi-Square and Log-Linear Analyses of Eyewitness Evidence (E), Physical Evidence (P) and Death Penalty Scruples (DP) in Relation to Verdict (V).

Factor	*df*	χ^2	*p*
	χ^2 Main Effect Analyses on Verdict		
Eyewitness Evidence	2	7.57	.02
Physical Evidence	2	11.02	.004
Death Penalty Scruples	2	.08	.96
	Model		
Marginals + (V × P)	44	37.01	.76
Marginals + (V × P) + (V × E)	42	29.19	.93
Residual from dropping (V × E)	2	7.82	> .025
Marginals + (V × P) + (V × E) + (V × DP)	40	29.11	.90
Residual from dropping (V × DP)	2	.08	.96
	Model of Best Fit		
Marginals + (V × P) + (V × E)	42	29.19	.93

ship between death-penalty scruples and verdicts found in Experiment 1 was replicated with the new measure of scruples, as were the weak relationships found in Experiment 1 between death-penalty scruples and the subjects' gender ($\chi^2(2) = 7.84$, $p < .025$) and political party ($\chi^2(4) = 8.48$, $p = .075$). Again as in Experiment 1, neither the subjects' gender ($\chi^2(1) = .04$) nor their political party affiliations ($\chi^2(2) = 3.20$, $p = .20$) was related to their verdict decisions.

Table 5 Percent Guilty Verdicts Reflecting Log-Linear Model of Best Fit from Table 4

	Physical Evidence		
Eyewitness Evidence	*Weak*	*Moderate*	*Strong*
Weak	45.0%	68.4%[1]	84.0%[1]
Moderate	50.0%	50.0%	70.0%
Strong	70.0%	75.0%	95.0%

[1]For these cells, n = 19. For others, n = 20.

GENERAL DISCUSSION

Taken together, these findings question the soundness of the conclusions reached by Haney (1984), Cowan et al. (1984), and Thompson et al. (1984) regarding the potential of the *Witherspoon* selection procedures to bias juries against defendants.

In examining the relationship between death-penalty attitudes (death-qualified versus *Witherspoon*-excludable jurors) and verdicts, the present experiments did not find any support whatsoever for these past studies, nor for the California Supreme Court opinion in *Hovey*. And whereas death-penalty scruples appeared to be related to the subjects' gender and political party, neither scruples nor gender nor political party predicted verdicts. Instead, we found that our subjects' verdict decisions were determined, virtually exclusively, by the strength of the evidence presented in the case, exactly as is legally prescribed. Furthermore, we did not find that death-penalty attitudes were more likely to predict verdicts in cases with weak evidence. On the other hand, the sentences the subjects chose to impose could be predicted by the degree of heinousness of the crime and, even more so, by the subjects' death-penalty attitudes as measured by the simulated *voir dire*.

These findings can be viewed from at least two different perspectives. One addresses the potential explanations for the different pattern of results produced by our procedures compared to those from previous studies. The second explores the implications of our findings, in the light of previous studies and consequent judicial decisions, for the legal issues surrounding the death-qualification procedures.

Explanations for Discrepancies Across Studies

Measurement Issues. One explanation for the differences between the past and present results is the possibility that our attitude measures failed adequately to assess the death-penalty scruples (or some related construct) and, therefore, their inability to predict verdicts was simply a result of a poorly constructed questionnaire. This explanation is untenable: Our measure of the death-penalty attitudes was highly predictive of sentencing decisions in both studies (only one of the 45 subjects that our procedures classified as being strongly against the death penalty chose the death penalty after deciding the defendant was guilty of first degree murder). Not only does this finding indicate a high degree of validity of the scruples measure, but it makes intuitive sense in terms of how sentencing decisions are made in actual capital trials. Most significantly, this finding strongly supports the rationale for and practice of the removal of strongly anti-death penalty veniremen from capital juries.

One might still argue that our procedures for measuring scruples, despite their ability to predict sentencing decisions, were somehow less "meaningful" than those used in prior research. Scruples may only predict verdicts, this tortuous argument goes, when "correctly" measured. However, the argument can be abruptly halted by a simple (methodo)logical observation: If the relationship between scruples and verdicts is that hypersensitive to the instrument and procedures used to assess scruples, then the true *legal* relevance of any social-science finding concerning verdicts can by definition be achieved only by measuring scruples in experiments *exactly* as they are determined by actual court-room *voir dire* procedures. As far as we know, this step has never been taken in a research study and it is unlikely that the extensive questioning of veniremen during the *Witherspoon voir dire* can be ade-

quately simulated. For example, one questionnaire used to help select jurors in a recent capital case in California (*People v. Davis*, 1995) included 23 different questions about death-penalty attitudes. Furthermore, returning to the argument critical of our scruples-measuring instrument, it would seem that our measurement procedures more closely approximate those used in real court cases than do those used in prior research. Unlike past studies, one of our scruples criteria was based on aggregating responses to a number of different, though related, death-penalty attitude items.

A different methodological line of thought would focus on the measurement of verdict decisions across different experiments. In Cowan et al. (1984), for example, subjects were asked to indicate whether they thought the defendant was guilty of 1st degree murder, 2nd degree murder, manslaughter, or not guilty. There were no differences in the rate at which death-qualified and excludable subjects chose either 1st (7.8% and 3.3%, respectively) or 2nd degree murder (21.3% and 23.3%, respectively) in pre-deliberation decisions (the same was true for post-deliberation decisions). The differences between death-qualified and excludable subjects emerged in the rates at which they chose the manslaughter and not guilty alternatives. In our studies, subjects simply indicated whether they felt the defendant was guilty of 1st degree murder or not. It is possible that scruples have no effect on the rate at which more severe verdicts are chosen. If so, it could be said that our results actually replicate a part of the Cowan et al. findings, namely, the failure of scruples to affect the rate of 1st degree murder decisions. This view of the data leads to the intriguing conclusion that attitudes toward the death penalty affect only the rate at which subjects pick verdicts that cannot possibly result in the death penalty. Are the subjects sophists? Are the jurors?

Case Differences. Another line of reasoning would pursue the idea that the differential presentation of the defendant's behavior across experiments, rather than the strength of the evidence against the defendant, is what matters. In Cowan et al., the issue that the simulated jurors had to decide was whether the defendant intended to harm the victim or did so accidentally, in an escalating sequence of aggressive encounters. In the present studies, the intention of the criminal to cause harm was clear. Simulated jurors had to decide whether the evidence was strong enough to tie the accused to the crime. Scruples may play a part only when the motives of the defendant are at issue, but not when the evidence is strong enough to convince the simulated jurors that the person being tried really committed the crime. From this viewpoint, one would hope that future studies vary even more features of the cases than the strength of evidence and heinousness. It seems especially important to investigate the various features of the type of cases in which the prosecution modally seeks a 1st degree murder conviction. Scruples may affect verdict choices, but only in cases in which the prosecution would *not* seek the death-penalty (such as cases used by Cowan et al., in which the link between the defendant's intention and the crime is so weak that even a charge of 1st degree murder would be unlikely). In other words, scruples may predict verdicts only in cases in which the question of juror death-qualification would be immaterial. More juror sophistry?

Theoretical (Social-Psychological) Issues. A still different approach to explaining our findings might focus on factors that have been shown to affect the attitude-behavior consistency, for example, the amount of direct experience on which the attitude is based (Fazio & Zanna,

1978) or the degree of presence of factors that might constrain behaviors that would otherwise reflect attitudes (Warner & DeFleur, 1969).

In the former case, one might argue that our college-age subjects were less likely than other registered voters to have had a direct experience with verdict and death-penalty decisions and therefore the relationship between the verdicts (behavior) and attitudes was not as strong as in previous research. But it is entirely unclear why college students who are taking philosophy, political science, economics, and sociology classes should not have more experience with death-penalty issues than do the average registered voters, especially those (as in the Moran-Comfort 1986 study) whose exposure to the local news media coverage of various felonies is so minimal that they manage to get impaneled (i.e., are not excused during *voir dire* for having already formed an opinion about the guilt of the defendant).

In the latter case, if the stimulus-constraint explanation were correct, one might have expected to find a stronger effect of scruples on verdicts in the weaker-evidence conditions of both experiments. This was not the case. The scruples measures were equally unable to predict verdicts at all evidence (and heinousness) levels. On the other hand, evidence by itself may not be the crucial constraining factor that prevents prior attitudes from affecting behavior in jury trials. Strong (and authoritatively delivered) judicial instructions to base a decision only on the evidence presented in court may be sufficient to cause jurors to set aside their opinions about the death penalty when contemplating the defendant's guilt. In the present research, we made a special point of emphasizing to all subjects the need to reach decisions solely on the basis of the evidence. It is unclear whether other researchers did the same. If this explanation is correct, it is imperative that future research make every effort to simulate, as closely as possible, not only the content, but the aura of authority and motivational force of CALJIC (or its equivalent) jury instructions. Our observation of *voir dire* in several capital cases suggests that judges and attorneys spend a considerable amount of time before, during, and after *voir dire* explaining that the law requires the verdict and sentence decisions to be independent.

Implications for the Legal System

The second perspective from which the present research can be viewed is the challenge it poses to the persistent attempts by social scientists (a relatively small group of interested researchers, actually) to influence legal thinking and judicial policy on death-qualification issues by reference to the previous simulation studies, the findings of which the present results directly contradict. We believe that their efforts are, and have been even prior to our studies, grossly misguided and misleading.

Regardless of the reasons for the differences in the results between the present and previous studies, the key question that plagues simulation research concerns external validity. As was mentioned in the introduction, we undertook this research within the simulation paradigm in order to make a point about the external validity of the bulk of past findings. In particular, although some previous studies have used broader subject sampling, they have been manifestly narrow in their sampling of case facts. Researchers who rely exclusively on this methodology (we certainly have not, in our other work on legal decision-making) usually defend their simulations as "more realistic" or as being more externally/ecologically valid

than other simulations that produce conflicting findings. Such defensive claims, however, are often logically groundless.

For example, some of the recent studies cited here used videotaped materials, rather than written summaries or questionnaires (Cowan et al., 1984; Thompson et al., 1984; Haney, 1984). We assume that this was done because of the authors' intuition that a videotaped presentation of events is somehow better at simulating reality than something written (or, more cynically, a part of a more convincing persuasion package for the courts). Without an empirical substantiation of this claim, it is equally reasonable to argue that in the age of television the subjects are totally desensitized to videotaped presentations of any kind and therefore give make-believe responses to them. Until evidence is painstakingly obtained that speaks directly to the external validity of findings across divergent media of stimulus presentations for specific populations, one is left not knowing which method best simulates the domain to which one wishes to generalize.

As researchers in personality have known for some time, the weakest form of validity is face validity. To infer that a procedure simulates basic processes because the procedure has some distant surface similarity to the generalization domain is simply incorrect. To infer that a conclusion has general value because several different methods yield similar results is certainly better logic, but its correctness depends on just how different the various methods really are and on the minute details of the requirement that the results be "consistent" across methods. Our results suggest that the method-specific variance may limit the generality of the relationship between the death-penalty attitudes and verdict decisions.

Besides, what is the significance of increased experimental realism for the issue of death-qualified juries? Trial procedures in a real court-of-law create a complex, intimidating, and powerful environment. It is a basic social-psychological principle that as situational influences increase, individual behavioral differences decrease (Mischel, 1976). It follows that the more realistic simulations would *suppress* the effects of individual attitudes on verdict decisions and increase the role of situational (i.e., case) factors such as evidence, to an even greater extent than the present—not very realistic—studies showed.

An additional point about external validity concerns the fact that the death-qualification portion of *voir dire* in capital cases comprises only a very small part of the entire pre-trial selection procedure. The majority of the questions normally asked of veniremen has been virtually ignored in the entire body of research on death-penalty scruples. Not only may prosecuting and defense attorneys use peremptory challenges to excuse veniremen for many reasons other than scruples, but judges typically eliminate a very large proportion of veniremen for cause and hardship. It is conceivable that in real trials, the small percentage of veniremen who would be excused on *Witherspoon* grounds (only 10–15% in our and prior studies), would be eliminated anyway on issues peripheral to the death-penalty attitudes (e.g., pre-trial knowledge of the case, attitudes about crime and criminals, past personal experience with violent crime, etc.).

For example, even though the O. J. Simpson trial is not a capital case, the questionnaire filled out by the potential jurors *prior* to *voir dire* consisted of 61 pages and contained 294 items, many of which had numerous parts with probing "Why do you believe that?" and "If yes, please explain" questions. Nor is this atypical. A questionnaire for veniremen in a recent, virtually unpublicized capital case in California contained over 140 different multi-

part questions. Clearly, the problem of jury bias that is addressed by the death-qualification research may be too circumscribed to warrant the attention it has received.

Judging from face value (an intellectual stance we normally do not recommend, but must now adopt in view of other authors' inferences from their research), the present experiments fully support the current use of death-qualified juries in capital cases. This is because our research, within its framework, unambiguously demonstrates (a) the overriding effect of the strength of evidence on verdict decisions, and (b) the virtually nonexistent effect of the *Witherspoon*-exclusion part of *voir dire* on the subjects' verdicts. To critics who might argue that such conclusions follow from "null" results, we respond that the very use of different scruples measures (an almost unprecedented feat in this area of research), the fact that our different scruples measures all successfully predicted sentencing behavior, and the fact that the effects of other variables were captured with the same measures, accumulate to a very substantial and consistent pattern of findings. The pattern becomes more significant when it is considered together with the conceptual and methodological critique of prior work.

In our view, there is no logical way to circumvent the paramount importance of thoroughly testing the external validity of research findings before applying them to the legal—or any other—system. The high-quality research evidence (in terms of methodological sophistication, data-collection thoroughness, ideological impartiality, etc.) simply does not exist to warrant a fundamental change in an established procedure, such as a shift to a dual-jury system for the two phases of capital-crime trials. We similarly deplore the dearth of evidence (i.e., Haney, 1984) on the basis of which the California Supreme Court mandated the costly and time-consuming sequestration of jurors during the death-qualification *voir dire*.

Monahan and Walker (1985, 1986) have observed that the U.S. legal system is gradually becoming more receptive to social science data. This is objectively an unfortunate development, unless the researchers seriously strive to design studies that may acquire "probative value" (Walker & Monahan, 1988). A social-science finding, by definition, cannot and should not have probative value without possessing a high degree of external validity. And it is extremely difficult for vociferous traditions in social science research—that rely almost exclusively on simulation methods that are often applied in a facile manner, and seem oblivious to their pitfalls—to acquire overnight the discipline and rigor that for example, the archival methodology requires.

CONCLUSIONS AND *WITT*

Since *Witherspoon*, the Supreme Court has consistently resisted social scientists' attempts to use the results from simulations to influence the process by which death-qualified juries are selected. The fifteen studies introduced in *Lockhart v. McCree* (1986) were individually criticized by the Court and the brief filed by the American Psychological association on that occasion (Bersoff & Ogden, 1987) fared no better. In fact, even before deciding *McCree*, the Court considerably modified the *Witherspoon* standard in *Wainwright v. Witt* (1985). By citing the language from *Adams v. Texas* (1980), that a potential juror may be excused for cause if his views would "prevent or substantially impair the performance of his duties as a juror in accordance with his instructions and oath," the *Witt* decision—the current death-qualifica-

tion standard—dispensed with the restrictive terminology of *Witherspoon* and essentially gave judges far broader discretion to exclude potential jurors.

Witt was predictably attacked on legal grounds (e.g., Callans, 1985; Meltzer, 1985; Rosenson, 1986; Thompson, 1989), as was the Court's related reasoning in and interpretation of *Adams, McCree*, and *Witherspoon* (e.g., Ellsworth, 1991; Faigman, 1991; Thompson, 1989). In some cases, the protests came from barely concealed ideological-*cum*-high-moral positions; occasionally they bordered on discourtesy to the Court, seemingly borne out of impotence to influence its decisions.

With Elliott (1991a, 1991b; 1992; Elliott & Robinson, 1991), we feel—and the data and methodological review presented here support us—that the social scientists' frustration has been well-earned through the inappropriateness of much research regarding death qualification. The Supreme Court has simply continued to indicate its disapproval of the quality of social-science advice it has received since *Witherspoon* invited it.

References

Adams v. Texas 448 U.S. 38 (1980).

Bersoff, D. N., & Ogden, D. W. (1987). In the Supreme Court of the United States: *Lockhart v. McCree. American Psychologist*, **42**, 59–68.

Bronson, E. (1970). On the conviction proneness and representativeness of the death qualified jury: An empirical study of Colorado veniremen. *University of Colorado Law Review*, **42**, 1–32.

Callans, P. J. (1985). Sixth Amendment—Assembling a jury willing to impose the death penalty: A new disregard for a capital defendant's rights. *Journal of Criminal Law & Criminology*, **76**, 1027–1050.

Cowan, C., Thompson, W. & Ellsworth, P. (1984). The effects of death qualification on juror's predisposition to convict and on the quality of deliberation. *Law and Human Behavior*, **8**, 53–60.

Ebbesen, E. B., & Konečni, V. J. (1980). On the external validity of decision-making research: What do we know about decisions in the real world? In T. S. Wallsten (Ed.) *Cognitive processes in choice and decision behavior*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Ebbesen, E. B., & Konečni, V. J. (1985). Criticisms of the criminal justice system: A decision-making analysis. *Behavioral Sciences and the Law*, **3**, 177–194.

Elliott, R. (1991a). Social science data and the APA: The *Lockhart* brief as a case in point. *Law and Human Behavior*, **15**, 59–76.

Elliott, R. (1991b). Response to Ellsworth. *Law and Human Behavior*, **15**, 91–94.

Elliott, R. (1992). On the alleged prosecution-proneness of death qualified jurors and juries. In P. Suedfeld & P. Tetlock (Eds.) *Psychology and social policy*. Washington, D.C.: Hemisphere.

Elliott, R., & Robinson, R. J. (1991). Death penalty attitudes and the tendency to convict or acquit: Some data. *Law and Human Behavior*, **15**, 389–404.

Ellsworth, P. C. (1991). To tell what we know or wait for Godot? *Law and Human Behavior*, **15**, 77–90.

Faigman, D. L. (1991). "Normative Constitutional fact-finding": Exploring the empirical component of Constitutional interpretation. *University of Pennsylvania Law Review*, **139**, 541–613.

Fazio, R. H., & Zanna, M. P. (1978). Attitudinal qualities relating to the strength of the attitude-behavior relationship. *Journal of Experimental Social Psychology*, **14**, 398–408.

Goldberg, F. (1970). Toward expansion of Witherspoon: Capital scruples, jury bias, and the use of psychological data to raise presumptions in the law. *Harvard Civil Rights–Civil Liberties Law Review*, **5**, 53–69.

Haney, C. (1984). On selection of capital juries: The biasing effects of the death qualifying process. *Law and Human Behavior*, **8**, 121–132.

Hovey v. California 28 Cal.3d 1 (1980).

Jurow, G. L. (1971). New data on the effects of a death qualified jury on the guilt determination process. *Harvard Law Review*, **84**, 567–611.

Konečni, V. J., & Ebbesen, E. B. (1979). External validity of research in legal psychology. *Law and Human Behavior*, **3**, 39–70.

Konečni, V. J., & Ebbesen, E. B. (1981). A critique of theory and method in social-psychological approaches to legal issues. In B. D. Sales (Ed.) *Perspectives in law and psychology, Volume II: The trial process*. New York: Plenum.

Konečni, V. J., & Ebbesen, E. B. (1982). Social psychology and the law: The choice of research problems, settings and methodology. In V. J. Konecni and E. B. Ebbesen (Eds.) *The criminal justice system: A social-psychological analysis*. San Francisco: W. H. Freeman & Co.

Konečni, V. J., & Ebbesen, E. B. (1984). The mythology of legal decision-making. *International Journal of Law and Psychiatry*, **7**, 5–18.

Konečni, V. J., & Ebbesen, E. B. (1992). Methodological issues in research on legal decision-making with special reference to experimental simulations. In F. Lösel, D. Bender, & T. Bliesener (Eds.) *Psychology and Law*. Berlin/New York: W. de Gruyter.

Lacayo, R. (1986). shaking the judicial perch. *Time*, **128**, September 15, 76.

Lockhart v. McCree, 106 S. Ct. 1758 (1986).

Meltzer, C. L. (1985). The battle over the standards for death-qualifying juries: Defendants lose another round. *Chicago Kent Law Review*, **61**, 611–628.

Mischel, W. (1976). The interaction of personality and the situation. In D. Magnusson and N. S. Miller (Eds.), *Personality at the crossroads: Current issues in interactional psychology*. Hillsdale, NJ: Earlbaum.

Monahan, J., & Walker, L. (1985). *Social science in law: Cases and materials*. Mineola, NY: Foundation Press.

Monahan, J., & Walker, L. (1986). Social authority: Obtaining, evaluating, and establishing social science in law. *University of Pennsylvania Law Review*, **134**, 477–517.

Moran, G., & Comfort, J. (1986). Neither "tentative" nor "fragmentary": Verdict preference of impaneled felony jurors as a function of attitude toward capital punishment. *Journal of Applied Psychology*, **71**, 146–155.

People v. Richard Allen Davis, Sonoma County, California (1985). Unpublished jury questionnaire.

Rosenson, V. T. (1986). *Wainwright v. Witt*: The Court casts a false light backward. *Boston University Law Review*, **66**, 311–343.

Tesser, A. (1978). Self-generated attitude change. *Advances in Experimental Social Psychology*, **11**, 289–338.

Thompson, W. C. (1989). Death qualification after *Wainwright v. Witt* and *Lockhart v. McCree*. *Law and Human Behavior*, **13**, 185–215.

Thompson, W. C., Cowan, C., Ellsworth, P., & Harrington, J. (1984). Death penalty attitudes and conviction proneness: The translation of attitudes into verdicts. *Law and Human Behavior*, **8**, 95–100.

Wainwright v. Witt 105 S. Ct. 844 (1985).

Walker, L., & Monahan, J. (1988). Social facts: Scientific methodology as legal precedent. *California Law Review*, **76**, 877–896.

Wilson, W. (1964). *Belief in capital punishment and jury performance*. Unpublished manuscript, University of Texas.

Witherspoon v. Illinois 391 U.S. 510 (1968).

Zeisel, H. (1968). *Some data on juror attitudes toward capital punishment*. University of Chicago Law school: Center for studies in criminal justice.

Notes

1. The authors wish to thank Lynda Tyll for her valuable participation on this project. Reprint requests may be sent to Vladimir J. Konečni or Ebbe B. Ebbesen, Department of Psychology 0109, University of California, San Diego, La Jolla, CA 92093.
2. Our data were collected before it became clear that the case of *Wainwright v. Witt* (1985) would set the new standard for the qualification of juries in capital-crime cases. We feel that the value of the research presented here is by no means diminished by the fact that *Witt* replaced *Witherspoon* as the standard for the selection of juries in capital cases. Rather, one can view the implications of the present body of data and methodological criticism as indicative of the type of skepticism that the Court may have covertly expressed in the *Witt* opinion regarding the merit and applicability of the tendentious social-science research that followed *Witherspoon*. See also the *Conclusions* section of this article with regard to *Witt*.
3. Chief Justice Rose Bird and two other California Supreme Court Justices, J. Grodin and C. Reynoso, failed to be re-elected (under confirmation law) by the people of California in 1986. It is generally agreed that this occurred because of the Court's frequent reversal of death-penalty verdicts in a State where 80% of the electorate favors capital punishment. Fifty-two of 55 death-penalty decisions heard by the Court during the tenure of these Justices were reversed (Lacayo, 1986).

4. This procedure is, of course, far removed from the manner in which attorneys and judges detect veniremen who are *Witherspoon* excludables. Questionnaires, whether consisting of one or two questions (as in the studies reported by Ellsworth and her colleagues, or by Haney) or consisting of nine items, as in the present research, may well not capture the decision rules used by the actual participants in the system to qualify jurors.
5. Log-linear odds analyses were used throughout this report rather than the analysis of variance because the main dependent variables were categorical, e.g., guilty vs. not guilty. Nevertheless, analyses of variance of the data completely supported the conclusions from the log-linear procedures.

APPENDIX 1
SAMPLE CASE SUMMARIES

CASE SUMMARY

EXPERIMENT 2

LOW EYEWITNESS EVIDENCE—LOW PHYSICAL EVIDENCE CONDITION

At 11:45 p.m. on November 2, 1985, a man entered a convenience store on El Cajon Boulevard, robbed the store, and shot the attendant to death. Approximately 5 minutes later, a police patrol car on its way to the scene, captured a man—the defendant, Robert Green—running down an alley near the scene of the crime. The police report stated that the store attendant was shot once through the chest and was dead at the scene. The cash register drawer was open and contained only pennies, nickles, dimes, and checks. The store owner estimated that there would have been approximately $250 to $300 in the register at the time of the robbery. The following is a summary of the main points in the case of The People versus Robert Green. The defendant is charged with first degree murder and has pleaded not guilty.

THE PROSECUTOR'S CASE IS AS FOLLOWS:

EYEWITNESS TESTIMONY: A customer in the store who hid behind some shelves and was not seen by the gunman, heard the entire episode. She claimed that the robber entered the store and demanded money. The store attendant, while complying, triggered the store alarm. When the alarm sounded, the robber shouted "Hey, no alarm!" and shot the attendant. The robber then put the money in a bag and ran out of the store. The witness further testified that as the gunman ran out of the store she was able to see that he was wearing faded blue levis and running shoes. Later, when the police brought the defendant back to the scene and asked him to speak, the witness said that the voice of the man in custody was the same as the voice of the gunman.

PHYSICAL EVIDENCE: A bag from the convenience store containing $283.00 was found in a trash dumpster about one block from where the defendant was apprehended. When the defendant was taken into custody he was wearing faded blue levis and running shoes.

THE DEFENSE ATTORNEY'S CASE IS AS FOLLOWS:

REBUTTAL TO EYEWITNESS TESTIMONY: In cross-examination it was pointed out that because the witness in the store was hiding and afraid to be seen by the gunman, she never actually saw anything and cannot identify the defendant as the robber except by the sound of his voice, which is subject to a great deal of human error. It was made clear that she is only speculating about what happened based on what she heard. From her position, crouched down behind the shelves, she was only able to catch a glimpse of the robber's clothes.

REBUTTAL TO PHYSICAL EVIDENCE: The defense argued that the levis and running shoes prove nothing since this is very common attire. The defendant testified that he was

innocently jogging through the alley when he was taken into custody by police. He said that he had not been in the convenience store at all that day and he had no knowledge of a murder or any money in a trash dumpster.

THE JUDGES CHARGE TO THE JURY IS AS FOLLOWS:

As jurors, it is very important to remember that the defendant is presumed to be innocent and the fact of being charged with a crime is not to be taken as evidence against him. You must find guilt beyond a reasonable doubt in order to convict. It is rarely possible to prove anything to an absolute certainty, but the evidence must be such that you, the jury, are willing to rely and act upon it.

CASE SUMMARY

EXPERIMENT 1

HIGH EVIDENCE—HIGH HEINOUSNESS CONDITION

At 11:45 p.m. on November 2, 1985, a man entered a convenience store on El Cajon Boulevard, pistol whipped the attendant repeatedly about the head and face, tied him up, forced to kneel, and shot him through the head, killing him instantly. The gunman then shot the body of the store attendant five more times, took the money from the cash register and ran from the store. Approximately 5 minutes later, a patrol car on its way to the scene, captured a man—the defendant, Robert Green—running down an alley 4 blocks from the scene of the crime.

THE PROSECUTOR'S CASE IS AS FOLLOWS: There were two customers, a man and a woman, in the rear of the store during the entire episode. Both crouched down behind some shelves and were not seen by the gunman. Although they were afraid to move and risk drawing attention to themselves, they could catch glimpses of the crime through cracks between the products on the shelves and were able to hear everything that happened. Their reports of the events of the crime were substantially in agreement: A man entered the store with a gun, beat the attendant brutally and repeatedly with the gun, tied him up, forced him to kneel, and shot him through the head. The gunman then fired five more bullets into the body of the attendant, opened the cash register, filled a bag with money, and ran out of the store. Both witnesses positively identified the defendant, Robert Green, as the man thy saw commit the crime. Both agree that the perpetrator was wearing light blue sweat pants, a blue levis denim jacket, and running shoes. When the defendant was taken into custody, he was wearing a t-shirt, light blue sweat pants, and running shoes. Additionally, when captured, he was in possession of $280.00 and could not account for where he obtained the money. Finally, the murder weapon, wrapped in a blue levis jacket was found in a trash dumpster in the alley where the defendant was captured. The jacket fit the defendant.

THE DEFENSE ATTORNEY'S CASE IS AS FOLLOWS: The witnesses' testimony is suspect because they could not see well enough from their distant and cramped vantage points

and would certainly have been too terrified for their own safety to accurately remember what they did see. On the stand, the defendant testified that he had won the money at the race track that afternoon, and that the levis jacket found in the dumpster was not his. The defense attorney pointed out that sweat pants and running shoes prove nothing since this is very common attire. Additionally, there were no readable fingerprints on the murder weapon. His client, Robert Green, was simply out jogging on the evening in question. When taken into custody, he knew nothing about a murder or robbery.

THE JUDGE'S CHARGE TO THE JURY IS AS FOLLOWS: As jurors, it is very important to remember that the defendant is presumed to be innocent and the fact of being charged with the crime is not to be taken as evidence against him. You must find guilt beyond a reasonable doubt in order to convict. It is rarely possible to prove anything to an absolute certainty, but the evidence must be such that you are willing to rely and act upon it.

APPENDIX 2
DEATH-PENALTY ATTITUDE QUESTIONNAIRES

DEATH PENALTY QUESTIONNAIRE

(EXPERIMENT 1)

PLEASE RESPOND TO THE STATEMENTS BELOW BY INDICATING AFTER EACH WHETHER YOU AGREE, ARE NOT SURE, OR DISAGREE WITH THE STATEMENT.

1. I entertain a conscientious opinion regarding the death penalty such that I could not make an impartial decision as to the defendant's guilt or innocence regardless of the evidence presented.
   _____AGREE _____NOT SURE _____DISAGREE

2. I entertain such conscientious opinions regarding the death penalty that I would automatically and without regard to the evidence refuse to find the alleged special circumstances to be true.
   _____AGREE _____NOT SURE _____DISAGREE

3. I entertain conscientious opinions regarding the death penalty such that if I found the defendant guilty of murder in the first degree, I would automatically and without regard for the evidence find the special circumstances to be true.
   _____AGREE _____NOT SURE _____DISAGREE

4. If I found the defendant guilty of murder in the first degree, and further found one of the special circumstances to be true, I would then feel compelled to automatically vote for the death penalty.
   _____AGREE _____NOT SURE _____DISAGREE

5. If I found the special circumstances allegation to be true, I would automatically and without regard for the evidence, vote for life imprisonment without the possibility of parole.
   _____AGREE _____NOT SURE _____DISAGREE

6. I am opposed to the use of the death penalty no matter what the crime.
   _____AGREE _____NOT SURE _____DISAGREE

7. When someone is found guilty of first degree murder, I believe the death penalty should always be the punishment used.
   _____AGREE _____NOT SURE _____DISAGREE

8. When some is found guilty of first degree murder, there might still be certain circumstances which would make the death penalty NOT appropriate.
   _____AGREE _____NOT SURE _____DISAGREE

9. Sometimes a crime is so brutal and terrible that the death penalty is the only appropriate punishment.
   _____AGREE _____NOT SURE _____DISAGREE

10. I believe the death penalty should be used more often than it is now.
    _____AGREE _____NOT SURE _____DISAGREE

DEATH PENALTY ATTITUDE QUESTIONNAIRE

(EXPERIMENT 2)

PLEASE RESPOND TO THE STATEMENTS BELOW BY INDICATING AFTER EACH WHETHER YOU AGREE, ARE NOT SURE, OR DISAGREE WITH THE STATEMENT.

1. When someone is found guilty of first degree murder, I believe the death penalty should ALWAYS be the punishment used.
   _____AGREE _____NOT SURE _____DISAGREE

2. When someone is found guilty of first degree murder, there might still be certain circumstances which would make the death penalty NOT appropriate.
   _____AGREE _____NOT SURE _____DISAGREE

3. Sometimes a crime is so brutal and terrible that the death penalty is the ONLY appropriate punishment.
   _____AGREE _____NOT SURE _____DISAGREE

4. I am opposed to the use of the death penalty no matter what the crime.
   _____AGREE _____NOT SURE _____DISAGREE

5. The death penalty should be used more often than it is now.
   _____AGREE _____NOT SURE _____DISAGREE

6. If I were a juror in an actual murder trial and the defendant were found guilty of first degree murder, I would NOT be able to vote for the death penalty.
   _____AGREE _____NOT SURE _____DISAGREE

7. If, on the next election ballot, a measure appears to make the death penalty illegal, I will vote against it (vote to keep the death penalty legal).
   _____AGREE _____NOT SURE _____DISAGREE

8. The state does NOT have the right to take someone's life no matter what the reason.
   _____AGREE _____NOT SURE _____DISAGREE

9. Life imprisonment without the possibility of parole is preferable to the death penalty as the maximum penalty allowed by law for capital crimes.
   _____AGREE _____NOT SURE _____DISAGREE

10. Anyone who intentionally kills another person deserves the same fate.
    _____AGREE _____NOT SURE _____DISAGREE