THE LEGAL AND METHODOLOGICAL IMPLICATIONS OF DEATH QUALIFICATION OPERATIONALIZATION

Matthew P. West, Emily F. Wood, Jena B. Casas, and Monica K. Miller
University of Nevada

Jury selection for death penalty cases is unique in that it includes a "death qualification" process in which prospective jurors are probed concerning their attitudes toward the death penalty and whether their attitudes would interfere with their performance as jurors. The current study assessed the relationship between two death qualification standards (i.e., the *Witt* standard and *Witherspoon-Morgan* standard) and mock jurors' endorsements of evidence and sentencing decisions. Results showed the *Witherspoon-Morgan* standard was significantly related to the dichotomous sentencing decision, whereas the *Witt* standard was not. Similarly, the *Witt* standard was significantly related to aggravator endorsement, whereas the *Witherspoon-Morgan* standard was not. These disparate relationships have implications for attorneys and judges in capital cases, as well as death penalty researchers.

Keywords: death qualification, capital punishment, juror decision-making, death penalty attitudes, jury selection

Jury selection for death penalty cases is unique in that it includes a "death qualification" process. During this process, prospective capital jurors are probed concerning their attitudes toward the death penalty and whether or not their attitudes would interfere with their ability to serve as impartial jurors. Prospective jurors who report their attitudes would affect their ability to serve as jurors can be excluded according to the predominant *Witt* death qualification standard (*Wainwright v. Witt*, 1985; see also *Lockhart v. McCree,* 1986). Prospective jurors who report they would always vote for a life sentence can be excluded according to the *Witherspoon* standard (*Witherspoon v. Illinois,* 1968); similarly, prospective jurors who report they would always vote for the death penalty are excludable according to the *Morgan* standard (sometimes referred to as "life qualification;" *Morgan v. Illinois,* 1992).

Author Note: Matthew P. West, Department of Criminal Justice, University of Nevada, Las Vegas; Emily F. Wood, Interdisciplinary Social Psychology Ph.D. Program, University of Nevada, Reno; Jena B. Casas, Department of Criminal Justice, University of Nevada, Reno; Monica K. Miller, Department of Criminal Justice Department and Interdisciplinary Social Psychology Ph.D. Program, University of Nevada, Reno

A preliminary version of this article was presented at the 2016 annual meeting of the American Psychology-Law Society, Atlanta, GA.

Correspondence concerning this article should be addressed to Matthew P. West, Department of Criminal Justice, University of Nevada, Las Vegas, NV 89154. E-mail: matthew.west@unlv.edu

Operational definitions of death qualification have varied across studies (e.g., Cowan Thompson, & Ellsworth, 1984; Summers, Hayward, & Miller, 2010)—including studies that did not directly examine death qualification. For example, Lynch and Haney (2009) excluded jurors based on a combination of *Witherspoon, Witt*, and *Morgan* standards; Lieberman, Shoemaker, and Krauss (2014) used a "modified" *Witherspoon* standard in their study 1, and a *Witt* standard in study 2; Summers et al. (2010) used a *Witt* standard (see also Patry & Penrod, 2013). The extent to which different death qualification standards might produce different results has not been fully investigated. Thus, the purpose of the current study was to assess the potentially differential relationships between death qualification standards and mock capital jurors' endorsements of evidence and sentencing decisions.

If there are differences between death qualification standards, there are critical legal and research implications. First, as Belt (1994) has suggested, capital juries will only "achieve real neutrality when the standards for assessing attitudes toward the death penalty are consistent" (p. 170). Jurors who are death qualified based on the criteria of one standard (e.g., *Witt* standard) might be more or less receptive to certain types of evidence or render different sentences than jurors death qualified according to a different standard (e.g., *Witherspoon* standard; however, see Butler & Moran, 2002). In turn, a capital defendant's risk of receiving the death penalty might vary according to a jurisdiction's particular application of death qualification standards (e.g., using *Witt* but not *Witherspoon* or *Morgan*). If a court solely applies one standard, the sentencing decision potentially could be overturned on constitutional grounds (i.e., impartial jury; 6th and 14th Amendments; Belt, 1994) if a defendant believes he would have received a more lenient sentence had the jury been death qualified using another standard (cf., *Uttecht v. Brown*, 2007).

Second, due to the inconsistencies in death qualification standards applied in death penalty research, this project could highlight important methodological considerations for future capital trial research. Moreover, addressing the potential impact of applying different death qualification standards provides a basis for more specific interpretations of past death penalty research (e.g., previously noted effects in past literature might be, at least in part, contingent upon the particular death qualification standard used). The subsequent sections will provide an overview of the modern capital trial format, past research on the influence of death qualification, and the Supreme Court cases in which death qualification standards have been promulgated.

CAPITAL TRIAL FORMAT

Today, the standard model for capital trials is a bifurcated process (*Gregg v. Georgia,* 1976; Palmer, 2014); that is, capital trials have two stages (i.e., guilt phase and penalty phase). In the guilt phase, jurors render the defendant guilty or not guilty. If the defendant is found guilty, jurors in the penalty phase then determine a sentence, either the death penalty or a life sentence (typically without the possibility of parole). At the penalty phase, jurors weigh aggravating and mitigating circumstances to arrive at a sentencing decision. [1] Aggravating circumstances (aggravators) make the defendant more deserving of the death penalty (e.g., the heinousness of the crime) and are determined by statute. Mitigating

circumstances (mitigators) make the defendant more deserving of a life sentence (e.g., the defendant cooperated with the authorities) and are partially determined by statute—jurors can consider virtually any case fact to be a mitigating factor (*Lockett v. Ohio,* 1978). In order to render a death penalty decision, jurors must find at least one aggravating circumstance present (*Zant v. Stephens,* 1983) and deem the aggravating circumstances sufficient to warrant the death penalty when considered with the mitigating circumstances (*Hurst v. Florida*, 2016; *Ring v. Arizona*, 2002). The present study simulates the sentencing phase of a capital trial.

DEATH QUALIFICATION

Jurors must be deemed "death qualified" to serve on a capital trial (see *Lockhart v. McCree,* 1986). The death qualification process is conducted to reveal potential jurors' attitudes toward the death penalty and to decide if the strength of their positions would conflict with their ability to be an impartial juror and follow the law. Death qualification standards have evolved over time. The standard(s) that are relied on or emphasized in capital *voir dire* might vary by jurisdiction.[2] Similarly, the standards applied in legal decision-making research have varied across studies. Past research has investigated the biasing effects of death qualification on perceptions of evidence, sentences, and verdicts, including process and eligibility effects, as discussed below.

Death Qualification Standards

In *Witherspoon v. Illinois* (1968), the Court articulated that prospective capital jurors could be excluded if: 1) "they would automatically vote against the imposition of capital punishment without regard to any evidence that might be developed at the trial of the case before them," or 2) "their attitude toward the death penalty would prevent them from making an impartial decision as to the defendant's guilt" (p. 522, n. 21). Almost two decades later, the Court noted how lower courts had been excluding jurors primarily using the first criterion of the *Witherspoon* standard—jurors who reported they would always render a life sentence were excluded (*Wainwright v. Witt*, 1985; see *Burns v. Estelle,* 1979). The Court then articulated what has become the predominant death qualification standard: The *Witt* standard. According to the *Witt* standard, prospective jurors can be excluded if they indicate their "views would prevent or substantially impair the performance of their duties in accordance with their instructions or their oaths" (*Wainwright v. Witt*, 1985, p. 424 n.5). Finally, in *Morgan v. Illinois* (1992), the Court drew on the rulings in *Wainwright v. Witt* (1985) and *Witherspoon v. Illinois* (1968), ruling that prospective jurors who would always vote for the death penalty may be excluded.[3]

Overall, in the death penalty studies published in the last decade, the majority have exclusively applied the *Witt* standard (e.g., Butler, 2007; Miller, Wood, & Chomos, 2014; Patry & Penrod, 2013; Summers et al., 2010). Some studies, however, have applied different standards (e.g., see Lieberman et al., 2014) or a combination of standards (e.g., Lynch & Haney, 2009). Therefore, this project also has methodological implications for future death penalty research.

Death Qualification and Decision-Making

The biasing effects of death qualification can be placed into two categories: Process effects and eligibility effects. The former, proposed and empirically examined by Haney (1984, 2005), posits that simply undergoing the death qualification process results in more punitive decisions. Participants in Haney's (1984) study who were exposed to death qualification deemed the defendant more culpable and were more likely to choose a death sentence than participants who had not been exposed to death qualification. Thus, the death qualification process independently results in more punitive decision-making.

Eligibility effects focus on how death qualified jurors differ from excludable jurors in their decision-making and their individual characteristics. Death qualified jurors are more likely to endorse aggravating circumstances (Butler & Moran, 2002, 2007a; Haney, Hurtado, & Vega, 1994); less likely to endorse mitigating circumstances (Butler & Moran, 2002, 2007a; Haney et al., 1994; Luginbuhl & Middendorf, 1988; Robinson, 1993); more likely to convict (Butler & Moran, 2007b, Cowan et al., 1984; Haney, 2005); and more likely to render a death sentence (Butler & Moran, 2002, 2007ab; Haney, 2005; see also Dillehay & Sandys, 1996; Sandys & Trahan, 2008). Importantly, death qualified jurors are more likely to convict and more likely to render a death sentence even after deliberation (Cowan et al., 1984). Death qualified jurors also differ from excludable jurors on myriad individual characteristics. For instance, death qualified jurors are more likely to be males and White (Butler & Moran, 2002; Summers et al., 2010; see also Sullivan, 2014; Swafford, 2011); be politically conservative (Butler & Moran, 2002, 2007a); hold more punitive attitudes toward criminals (Haney et al., 1994); and interpret the Bible literally (Summers et al., 2010). In sum, this evidence strongly indicates death qualified jurors are not representative of the general population (see Salgado, 2005; Thompson, 1989 for further discussion).

The last known study that reported the effects of multiple death qualification standards on aggravator and mitigator endorsement in the context of a simulated trial was over a decade ago (Butler & Moran, 2002). In that study, mock jurors death qualified according to the *Witt* standard were more likely to endorse aggravators, less likely to endorse mitigators, and more likely to render the death penalty than excludable jurors. The authors report in a footnote that the *Witherspoon-Morgan* standard was associated with aggravator and mitigator endorsement, but do not report posthoc comparisons or whether the *Witherspoon-Morgan* standard was associated with sentence.[4] Thus, although it appears clear that death qualification is generally related to punitive decision-making, it is unclear to what extent this relationship is present under different death qualification standards (see Butler & Moran, 2002; Haney et al., 1994; Sandys & Trahan, 2008; Rozelle, 2002).

CURRENT STUDY

The purpose of the current study was to address the gap in the literature concerning the differential relationships between various death qualification standards and jurors' sentencing decisions and endorsements of aggravators and mitigators. We sought to replicate Butler and Moran (2002), but also extend the literature by including another variable: The extent to which jurors deemed aggravating circumstances sufficient to warrant the death

penalty (see *Hurst v. Florida*, 2016; *Ring v. Arizona*, 2002). We operationalized two death qualification standards: A *Witt* standard (i.e., whether participants reported their attitudes toward the death penalty would affect their ability to be a capital juror) and a *Witherspoon-Morgan* standard (i.e., whether participants reported they would always render a life sentence or that they would always render a death sentence). This study was guided by the following hypothesis and research question:

> ***Hypothesis:*** Based on the findings of Butler and Moran (2002, 2007a), jurors who are death qualified according to the *Witt* standard will be more likely to endorse aggravating circumstances, less likely to endorse mitigating circumstances, more likely to deem aggravating circumstances sufficient for the death penalty, more likely to render the death penalty, and more likely to indicate certainty in a death sentence compared to jurors excludable according to the *Witt* standard.
>
> ***Research question*:** Will the pattern of results hypothesized above occur when jurors are death qualified according to the *Witherspoon-Morgan* standard instead of the *Witt* standard?

METHOD

Participants and Procedure

Participants were ($n = 457$) community members recruited via Amazon's Mechanical Turk who received $5.00. Participants were mostly male (56%) and ranged from 18 to 67 years, ($M = 35$; $Mdn = 33$). Participants were White (83%), Hispanic (5%), African American (6%), and Asian (6%). Eighteen percent of participants had previous experience serving as a juror. They were Catholic (15%), non-Catholic Christian (22%), atheist (21%), agnostic (20%), or indicated a belief in God without an affiliation (13%).

Participants first answered two death qualification questions. Next, participants read a summary of a sentencing phase of a capital trial. Participants rated the extent to which they agreed that each aggravator and mitigator were present in the case, rated the extent to which aggravators were sufficient to warrant the death penalty, reported their sentence recommendation, and indicated how certain they were in their sentence.

Materials

Death qualification questions. Death qualification was measured using two questions. The first question assessed participants' attitude toward the death penalty and asked, "What is your attitude toward the death penalty?" Participants were instructed to check one of the following statements indicating that they: 1) would always vote for the death penalty; 2) are in favor of the death penalty, but would not necessarily vote for it in every case; 3) have doubts about the death penalty, but would be able to find the defendant guilty and 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. This question was used to determine if participants were death qualified or excludable according to the *Witherspoon-Morgan* standard.

The second question assessed how participants would conduct themselves in a capital trial and asked, "Given your position regarding the death penalty, which of the following statements best describes how you would conduct yourself as a juror in a capital murder case?" Participants were instructed to check one of the following statements: 1) 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 2) 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.[5] This question was used to determine if participants were death qualified or excludable according to the *Witt* standard.

Case summary. Participants read an approximately 1,900-word summary based on a real murder case in North Carolina (*State v. Daniels,* 1994). The summary included case facts, closing arguments from the prosecution and the defense, and judge's instructions.

Endorsement of aggravators and mitigators. Endorsement of aggravators and mitigators was measured by asking participants to indicate the extent to which they agreed that each of the four aggravators and five mitigators were present in the case on a scale from 1 = *Strongly disagree* to 7 = *Strongly agree*. Participants did this for each aggravator (e.g., the capitol felony was especially heinous, atrocious, or cruel) and mitigator (e.g., the defendant is a good candidate for successful psychological rehabilitation; see North Carolina General Statutes, n.d.). Mean endorsement of aggravators and mean endorsement of mitigators variables were created by averaging scores on the aggravator items ($\alpha = 0.60$) and the mitigator items ($\alpha = 0.71$).

Aggravator sufficiency. Aggravator sufficiency was measured by asking participants to respond to the question, "Please indicate the extent to which you believe that the aggravating circumstance or circumstances found in issue one is or are sufficiently substantial to recommend the death penalty when considered with the mitigating circumstance or circumstances" on a scale from 1 = *Not substantial enough to recommend the death penalty* to 7 = *Completely substantial enough to recommend in favor of the death penalty.*

Sentencing decision. Sentencing decision was measured by asking participants to respond to the question, "What sentence do you recommend in this case?" by choosing "Death penalty" or "Life sentence without the possibility of parole."

Sentencing decision certainty. Sentencing decision certainty was measured by asking participants, "How certain are you that your sentencing decision is appropriate?" on a scale from 1 = *Very certain in a 'death penalty' decision* to 7 = *Very certain in a 'life sentence without the possibility of parole' decision.*

RESULTS

To address how the *Witt* standard and *Witherspoon-Morgan* standard were related to mean endorsement of aggravators, mean endorsement of mitigators, aggravator sufficiency, sentencing decision, and sentencing decision certainty, two sets of ANOVAs and two cross tabulations were conducted. The two sets of ANOVAs examined the relationships between both standards (independent variables) and the continuous dependent variables (i.e., mean endorsement of aggravators, mean endorsement of mitigators, aggravator sufficiency, and sentencing decision certainty); the two cross tabulations examined the relationship between both standards and the dichotomous sentencing decision. During initial data inspection, the sentencing decision certainty variable was found to be substantially skewed. Thus, a log transformation was applied following procedures suggested by Tabachnick and Fidell (2007) and Howell (2007). Following the transformation, skew was reduced to a more acceptable level.

Of the total sample, 2% (n=11) reported they would always support the death penalty; 40% (n=185) were in favor of the death penalty, but only under some circumstances; 42% (n=194) had doubts about the death penalty, but would support it if warranted; and 16% (n=74) reported they would never render the death penalty. Under the combined *Witherspoon v. Illinois* (1968) and *Morgan v. Illinois* (1992) standard (i.e., *Witherspoon-Morgan* standard), those that would always render the death penalty *or* would never render the death penalty were deemed excludable (n=85). This resulted in 372 *Witherspoon-Morgan* death qualified participants. Of the total participants, 17% (n=78) felt so strongly about the death penalty that they believed their views would affect their ability to be a capital juror; this rendered those participants excludable based on the *Witt* standard. This resulted in 379 *Witt* death qualified participants.

The first set of ANOVAs examined the relationship between the *Witt* standard (i.e., "*Witt* death qualified" compared to "*Witt* excludable" jurors) and sentencing decision certainty, aggravator sufficiency, and endorsements of aggravators and mitigators. The *Witt* standard was associated with aggravator endorsement ($F(1, 455) = 4.0$, $p = .046$, $\eta^2_{Partial} = 0.01$) but not mitigator endorsement. The *Witt* standard was significantly associated with aggravator sufficiency ($F(1, 455) = 9.59$, $p = .002$, $\eta^2_{Partial} = 0.02$) and sentencing decision certainty ($F(1, 455) = 15.85$, $p < .001$, $\eta^2_{Partial} = 0.03$). Posthoc comparisons using Bonferroni adjustments indicated *Witt* death qualified jurors were more likely to endorse aggravators ($M = 5.29$, $SE = 0.13$) than *Witt* excludable jurors ($M = 4.99$, $SE = 0.13$; $M_{Diff} = -0.29$, $p = .046$); more likely to find aggravators sufficient to warrant the death penalty ($M = 3.48$, $SE = 0.10$) than *Witt* excludable jurors ($M = 2.73$, $SE = 0.22$; $M_{Diff} = -0.75$, $p = .002$); and more likely to indicate certainty in a death penalty decision ($M = 0.34$, $SE = 0.02$) than *Witt* excludable jurors ($M = 0.20$, $SE = 0.03$; $M_{Diff} = -0.14$, $p < .001$). These results offer partial support for the hypothesis.

Finally, the association between the *Witt* standard and dichotomous sentencing decision was assessed using a cross tabulation and adjusted standardized residuals. The *Witt* standard was not significantly associated with sentence. This does not support the hypothesis.

The second set of ANOVAs examined the relationship between the *Witherspoon-Morgan* standard (i.e., "*Witherspoon-Morgan* death qualified" compared to "*Witherspoon-Morgan* excludable" jurors) and sentencing decision certainty, aggravator sufficiency, and endorsements of aggravators and mitigators. The *Witherspoon-Morgan* standard was not associated with aggravator endorsement or mitigator endorsement, but was significantly associated with aggravator sufficiency ($F(1, 455) = 12.44$, $p < .001$, $\eta^2_{Partial} = 0.03$) and sentencing decision certainty ($F(1, 455) = 18.53$, $p < .001$, $\eta^2_{Partial} = 0.04$). Posthoc comparisons indicated *Witherspoon-Morgan* death qualified jurors were more likely to find aggravators sufficient to warrant the death penalty ($M = 3.51$, $SE = 0.10$; $M_{Diff} = -0.83$, $p < .001$) compared to *Witherspoon-Morgan* excludable jurors ($M = 2.68$, $SE = 0.21$). Death qualified jurors were also more likely to indicate certainty in a death penalty decision ($M = 0.34$, $SE = 0.02$; $M_{Diff} = -0.15$, $p < .001$) than excludable jurors ($M = 0.20$, $SE = 0.03$).

Next, a cross tabulation was conducted to examine the association between the *Witherspoon-Morgan* standard and dichotomous sentencing decision. The *Witherspoon-Morgan* standard was significantly associated with sentence ($\chi^{(1)} = 4.27$, $\phi = 0.10$, $p = .04$). For *Witherspoon-Morgan* death qualified jurors (n=368), 85 rendered a death sentence (expected count = 78, $z = 2.1$, $p = .03$), and 283 rendered a life without possibility of parole sentence (expected count = 290, $z = -2.1$, $p = .03$). For *Witherspoon-Morgan* excludable jurors ($n = 85$), 11 rendered a death sentence (expected count = 18, $z = -2.1$, $p = .03$), and 74 rendered a life without possibility of parole sentence (expected count = 67, $z = 2.1$, $p = .03$).[6]

DISCUSSION

The purpose of the present study was to examine how death qualification (as operationalized by different standards) was related to aggravator endorsement, mitigator endorsement, aggravator sufficiency, sentencing decision, and sentencing decision certainty. We operationalized two death qualification standards: A *Witt* standard (based in *Wainwright v. Witt*, 1985) and a *Witherspoon-Morgan* standard (based in *Witherspoon v. Illinois*, 1968, and *Morgan v. Illinois*, 1992). It was hypothesized that *Witt* death qualified jurors would be more likely to endorse aggravators, less likely to support mitigators, more likely to deem aggravators sufficient for the death penalty, more likely to render the death penalty, and more likely to indicate certainty in a death penalty decision than *Witt* excludable jurors. The research question asked whether this hypothesis would apply when jurors were death qualified according to the *Witherspoon-Morgan* standard instead of the *Witt* standard. Results offered partial support for the hypothesis and indicated death qualification standards were differentially related to the dependent variables.

The differential relationships between death qualification standards were most notably exemplified in aggravator endorsement and the dichotomous sentencing decision. *Witt* death qualified jurors were more likely to endorse aggravating circumstances than *Witt* excludable jurors. Conversely, *Witherspoon-Morgan* death qualified jurors did not endorse aggravating circumstances differently than *Witherspoon-Morgan* excludable jurors. *Witherspoon-Morgan* death qualified jurors were more likely to render a death penalty decision and less likely to render a life sentence decision than *Witherspoon-Morgan* ex-

cludable jurors. However, there was no similar relationship found when using the *Witt* standard. This finding is particularly interesting because *both* standards were associated with sentencing decision certainty. It is possible that differences in sentencing decisions between *Witt* death qualified jurors and *Witt* excludable jurors can only be captured with a sensitive measure such as a certainty scale. In general, the association between the *Witt* standard and sentencing decisions might be less robust than the association between the *Witherspoon-Morgan* standard and sentencing decisions.

This study's findings in some ways run counter to previous death qualification literature. Butler and Moran (2002, 2007a) found that *Witt* death qualified jurors were more likely to endorse aggravating circumstance (replicated in the current study), less likely to endorse mitigating circumstances (not replicated), and more likely to render a death sentence than *Witt* excludable jurors (replicated with certainty scale but not dichotomous sentencing decision). Further, in their earlier article (i.e., Butler & Moran, 2002), they found that the *Witherspoon-Morgan* standard was associated with endorsements of aggravating circumstances and endorsements of mitigating circumstances (not replicated in the current study). Although the current study did not replicate all of these findings, it did expand on them: *Witt* and *Witherspoon-Morgan* death qualified jurors were more likely to deem aggravators sufficient to warrant the death penalty when considered with mitigators compared to *Witt* and *Witherspoon-Morgan* excludable jurors. No previous study has used this as a dependent variable, but this finding coalesces with previous literature suggesting that death qualified jurors are biased toward aggravating evidence (e.g., Haney, 2005). The finding that both standards were related to aggravator sufficiency but only the *Witt* standard was associated with aggravator endorsement might suggest there are nuances to the relationship between aggravator evaluations and death qualification according to standard. For example, *Witt* death qualified jurors might be more likely to perceive aggravators and therefore give them more subjective weight, whereas *Witherspoon-Morgan* death qualified jurors might not perceive aggravators differently than excludable jurors, but are predisposed to give aggravators more weight.

Critically, this study's method, materials, and participants differed from that of Butler and Moran's (2002, 2007a). One important difference between the current study and those studies is that, at least in the earlier Butler and Moran (2002) article, participants completed the study in a group setting; in contrast, the current study was conducted individually. Open-air group *voir dire* is associated with increased exclusions for cause based on *Witherspoon* (Nietzel & Dillehay, 1982). This might explain why, in the current study, 19% were *Witherspoon-Morgan* excludable, whereas in Butler and Moran's (2002) study, 30% were *Witherspoon-Morgan* excludable. Furthermore, we used different materials and measured aggravators and mitigators differently than Butler and Moran (2002, 2007a). Replication of the current study with a different sample, with different materials, and with individual and group death qualification is required to determine if these factors moderate the influence of death qualification.

Implications

This study has implications for the legal system and for death penalty researchers. First, there might be no way that death qualification can occur without resulting in bias, as both standards resulted in the inclusion of jurors who were biased against the defendant on most measures (see Haney, 2005). This study, along with numerous others (e.g., Butler, 2007; Butler & Moran, 2002; 2007ab; Cowan et al. 1984; Haney, 1984; 2005), have empirically demonstrated the biasing effect of death qualification on decision-making. However, the Court has essentially deemed this type of empirical evidence—even if the methodology is sound—irrelevant to the constitutionality of death qualification (*Lockhart v. McCree*, 1986; see also, *Uttecht v. Brown*, 2007). Thus, petitioning the Court on the basis of this literature might not be a promising route of reform. There might be alternatives to death qualification, though, that satisfy the need to impanel a jury that can follow the law without resulting in bias. One potential alternative proposed is a non-unanimous sentencing verdict (see Brief for Amicus Curiae American Psychological Association in Support of Respondent in *Lockhart v. McCree*, 1986; Tucker, 2012). By requiring a (super) majority instead of unanimity, jurors who might only consider a life sentence when a death sentence is legally appropriate could be included in jury deliberations, but would not inhibit a death penalty decision (Tucker, 2012). This is in need of empirical investigation.

There might be more practical reforms to increase consistency in the death qualification process. In practice, trial judges might interpret and apply death qualification standards in different ways. This could be why jurors who are technically excludable according to various standards—particularly those who are more likely to render a death sentence—often are not actually excluded (see Sandys & Trahan, 2008). Sandys and Trahan (2008) suggest that the attorneys and judge reach an agreement concerning *how* standards will be applied in a particular case beforehand. This agreed upon articulation could perhaps be used to develop a pretrial questionnaire. In this way, some prospective jurors could be excluded prior to *voir dire*, which Sandys and Trahan (2008) argue would allow for more thorough in-court death qualification. Results from this study showed differential effects of death qualification standards, but perhaps these discrepancies could be reduced if Sandys and Trahan's (2008) approach is implemented. Rozelle (2002) argues that death qualification should consist of a single question based on the *Witt* standard, similar to how the current study operationalized the *Witt* standard. Thus, the finding that the *Witt* standard was unrelated to the dichotomous sentencing decision could be preliminary evidence that using a strict, one question application of the *Witt* standard could avoid the biasing effect of death qualification on sentencing decisions. Even if Rozelle's (2002) application of the *Witt* standard did not avoid the death qualification bias, it would likely have the benefit of increasing consistency in how death qualification is conducted across jurisdictions because of its straightforward implementation (i.e., one question with a yes or no response).

Second, research using a capital trial context invariably includes a death qualification component, presumably in the interest of external validity. At minimum, researchers using a simulated capital case should clearly specify what standard they used and how participants were death qualified. The finding that death qualification standards had dispa-

rate associations with the dependent variables indicates it should be a carefully considered methodological component. Moreover, clearly specifying the death qualification component of a study allows for accurate replication. There is a definite possibility death qualification moderates the influence of other variables on sentences, verdicts, and aggravator and mitigator endorsement. In order to accurately replicate and subsequently interpret replications of research using a capital trial context, there needs to be methodological consistency. This is especially important if the empirical literature is to be used as a basis for influencing legal procedure and policy. Death penalty researchers might try running their analyses using different standards, or combination of standards, in order to determine whether the standard applied alters their results.

Similarly, death penalty researchers should consider including both a continuous measure of sentencing decision (e.g., a certainty scale) and a dichotomous measure of sentencing decision (i.e., death penalty or life in prison without possibility of parole). In the current study, *Witt* death qualified jurors were significantly more likely to be certain in a death penalty decision, but were not significantly more likely to choose a death sentence compared to *Witt* excludable jurors. A dichotomous sentencing measure might be more ecologically valid, but a sentence certainty scale might nevertheless indicate juror tendencies. Thus, including both types of measures might reveal more nuanced relationships. Ultimately, the researcher must consider the aims (along with the analysis plan, logistical limitations, etc.) of their specific project when determining how to operationalize sentencing decisions.

Limitations and Future Directions

This study has a few important limitations worth noting and addressing. Most critically, this study lacked consequentiality: Participants were only hypothetically rendering death and life sentences. The decision-making that results in a hypothetical death penalty decision might only resemble the decision-making that results in a real-world death penalty decision. However, past research suggests that simulated jury studies do not necessarily produce different results than studies of real-world jurors (Bornstein, 1999; Bornstein & McCabe, 2005; Bornstein et al., 2017; Kovera, 2017). Previous work on death qualification has used a sample of individuals called for jury service (e.g., Butler & Moran, 2002) and former capital jurors (e.g., Sandys & Trahan, 2008). Those types of samples arguably might be more representative of potential capital jurors, and thus this line of research could benefit from a replication of the current study using a sample of actual jurors (see Bornstein, 1999; Bornstein & McCabe, 2005; Bornstein et al., 2017; Kovera, 2017). Related, in real-world capital trials, attorneys sometimes attempt to "rehabilitate" excludable prospective jurors (e.g., the defense attorney might try to "save" a prospective juror from exclusion who initially reports they would never render a death penalty decision by urging them to consider circumstances they could hypothetically render a death penalty decision; Nietzel & Dillehay, 1982; Nietzel, Dillehay, & Himelein, 1987). Future research could examine this "rehabilitation" component of capital *voir dire*—particularly whether or not this component results in a prospective juror technically eligible for exclusion to nevertheless be death qualified. In addition, participants in the current study yielded individual sentences

and did not deliberate. Past research has found that death qualification bias persists through deliberation (Cowan et al., 1984), therefore, death qualified jurors would be expected to render similar sentencing decisions whether or not they underwent deliberation.

CONCLUSION

Death qualification is a unique requirement for capital jury service. Past research and the current study has suggested that death qualification results in biased endorsements of evidence and sentencing decisions (Butler & Moran, 2002, 2007a). This research demonstrated the differential relationships between death qualification standards and jurors' endorsement of evidence and sentencing decisions. Ultimately, the decisions that jurors in capital trials make have life altering consequences for capital defendants; thus, the way in which individuals are deemed eligible (i.e., the death qualification process) to serve on a jury is extremely important. Capital attorneys and judges might be able to increase consistency in death qualification by collectively determining how death qualification standards will be applied in a given case before the trial occurs (Sandys & Trahan, 2008). Additionally, future research should articulate *what* standard was applied, and, perhaps just as critically, *how* that standard was applied.

REFERENCES

Belt, J. C. (1994). Morgan v. Illinois: The right to balance capital sentencing juries as to their views on the death sentence is finally granted to defendants. *New Mexico Law Review*, *24*, 145-170. Retrieved from http://lawschool.unm.edu/nmlr/volumes/24/1/08_belt_morgan.pdf

Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior*, *23*(1), 75-91. doi: 10.1023/A:1022326807441

Bornstein, B. H., Golding, J. L., Neuschatz, J., Kimbrough, C., Reed, K., Magyarics, C., & Luecht, K. (2017). Mock juror sampling issues in jury simulation research: A meta-analysis. *Law and Human Behavior, 41*, 13-28. doi:10.1037/lhb0000223

Bornstein, B. H., & McCabe, S. G. (2005). Jurors of the absurd? The role of consequentiality in jury simulation research. *Florida State University Law Review, 32,* 443-465.

Burns v. Estelle, 592 F.2d 1297 (5th Cir. 1979).

Butler, B. (2007). The role of death qualification in capital trials involving juvenile defendants. *Journal of Applied Social Psychology, 37*(3), 549-560. doi:10.1111/j.1559-1816.2007.00174.x

Butler, B. M., & Moran, G. (2002). The role of death qualification in venirepersons' evaluations of aggravating and mitigating circumstances in capital trials. *Law and Human Behavior, 26*(2), 175-184. doi:10.1023/A:1014640025871

Butler, B., & Moran, G. (2007a). The impact of death qualification, belief in a just world, legal authoritarianism, and locus of control on venirepersons' evaluations of aggravating and mitigating circumstances in capital trials. *Behavioral Sciences & the Law, 25*, 57-68. doi: 10.1002/bsl.734

Butler, B., & Moran, G. (2007b). The role of death qualification and need for cognition in venirepersons' evaluations of expert scientific testimony in capital trials. *Behavioral Sciences & the Law, 25*(4), 561-571. doi:10.1002/bsl.758

Chappell v. State, 281 P.3d 1160 (Nev. 2009).

Cowan, C. L., Thompson, W. C., & Ellsworth, P. C. (1984). The effects of death qualification on jurors' predisposition to convict and on the quality of deliberation. *Law and Human Behavior, 8*(1-2), 53-79. doi:10.1007/BF01044351

Dillehay, R. C., & Sandys, M. R. (1996). Life under Wainwright v. Witt: Juror dispositions and death qualification. *Law and Human Behavior, 20*(2), 147-165. doi:10.1007/BF01499352

Gregg v. Georgia, 428 U.S. 153 (1976).
Haney, C. (1984). On the selection of capital juries: The biasing effects of the death-qualification process. *Law and Human Behavior, 8*(1-2), 121-132. doi:10.1007/BF01044355
Haney, C. (2005). *Death by design: Capital punishment as social psychological system*. New York: Oxford University Press.
Haney, C., Hurtado, A., & Vega, L. (1994). 'Modern' death qualification: New data on its biasing effects. *Law and Human Behavior*, *18*(6), 619-633. doi:10.1007/BF01499328
Howell, D. C. (2007). *Statistical methods for psychology* (6th ed.). Belmont, CA: Thomson Wadsworth.
Hurst v. Florida, 136 S. Ct. 616 (2016).
Jurek v. Texas, 428 U.S. 262 (1976).
Kovera, M. B. (2017). *The psychology of juries*. Washington, DC: American Psychological Association.
Lieberman, J. D., Shoemaker, J., & Krauss, D. A. (2014). The effects of mortality salience and evidence strength on death penalty sentencing decisions. *Psychology, Crime & Law, 20*(3), 199-221. doi:10.1080/1068316X.2013.770853
Lockett v. Ohio, 438 U.S. 586 (1978).
Lockhart v. McCree, 476 U.S. 162 (1986).
Luginbuhl, J., & Middendorf, K. (1988). Death penalty beliefs and jurors' responses to aggravating and mitigating circumstances in capital trials. *Law and Human Behavior, 12*(3), 263-281. doi:10.1007/BF01044384
Lynch, M., & Haney, C. (2009). Capital jury deliberation: Effects on death sentencing, comprehension, and discrimination. *Law and Human Behavior, 33*(6), 481-496. doi:10.1007/s10979-008-9168-2
McNally, K. M. (1985). The death penalty. *The Advocate*, *7, 12-18.*
Miller, M., Wood, S. M., & Chomos, J. C. (2014). Relationships between support for the death penalty and cognitive processing: A comparison of students and community members. *Criminal Justice and Behavior, 41*(6), 732-750. doi:10.1177/0093854813509369
Morgan v. Illinois, 504 U.S. 719 (1992).
Nietzel, M. T., & Dillehay, R. C. (1982). The effects of variations in voir dire procedures in capital murder trials. *Law and Human Behavior, 6*(1), 1-13. doi:10.1007/BF01049309
Nietzel, M. T., Dillehay, R. C., & Himelein, M. J. (1987). Effects of voir dire variations in capital trials: A replication and extension. *Behavioral Sciences & the Law, 5*(4), 467-477. doi:10.1002/bsl.2370050410
North Carolina Capital Punishment (1910). N.C. Gen. Stat. § 15-100. Retrieved March 8, 2015 from http://www.ncga.state.nc.us/enactedlegislation/statutes/html/bysection/chapter_15a/gs_15a-2000.html
Palmer, L. J., Jr. (2014). *The death penalty in the United States: A complete guide to federal and state laws*. Jefferson, NC: McFarland and Company.
Patry, M. W., & Penrod, S. D. (2013). Death penalty decisions: Instruction comprehension, attitudes, and decision mediators. *Journal of Forensic Psychology Practice, 13*(3), 204-244. doi:10.1080/15228932.2013.795816
Ring v. Arizona, 536 U.S. 584 (2002).
Robinson, R. J. (1993). What does "unwilling" to impose the death penalty mean anyway? Another look at excludable jurors. *Law and Human Behavior, 17*(4), 471-477. doi:10.1007/BF01044379
Rozelle, S. D. (2002). The utility of Witt: Understanding the language of death qualification. *Baylor Law Review, 54*, 677-699.
Salgado, R. (2005). Tribunals organized to convict: Searching for a lesser evil in the capital juror death-qualification process in United States v. Green. *Brigham Young University Law Review*, 519-552. Retrieved from http://digitalcommons.law.byu.edu/cgi/viewcontent.cgi?article=2234&context=lawreview
Sandys, M., & Trahan, A. (2008). Life qualification, automatic death penalty voter status, and juror decision making in capital cases. *The Justice System Journal, 29*(3), 385-395.
State v. Daniels, 446 S.E.2d 298, (1994).
Sullivan, T. J. (2014). The demographic dilemma in death qualification of capital jurors. *Wake Forest Law Review, 49*, 1107-1172.

Summers, A., Hayward, R. D., & Miller, M. K. (2010). Death qualification as systematic exclusion of jurors with certain religious and other characteristics. *Journal of Applied Social Psychology, 40*(12), 3218-3234. doi:10.1111/j.1559-1816.2010.00698.x

Swafford, A. T. (2011). Qualified support: Death qualification, equal protection, and race. *American Journal of Criminal Law, 39*(1), 147-174.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn and Bacon.

Tucker, J. (2012). *An alternative to death-qualification: The non-unanimous penalty jury* (Doctoral dissertation). Retrieved from Yale Law School Scholarship Repository.

Thompson, W. C. (1989). Death qualification after Wainwright v. Witt and Lockhart v. McCree. *Law and Human Behavior*, *13*(2), 185-215. doi:10.1007/BF01055923

Uttecht v. Brown, 127 S. Ct. 2218 (2007).

Wainwright v. Witt, 469 U.S. 412 (1985).

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

Zant v. Stephens, 462 U.S. 862 (1983).

Date Received: 10/2016
Date Accepted: 03/2017

Suggested citation: West, M. P., Wood, E. F., Casas, J. B., & Miller, M. K. (2017). The legal and methodological implications of death qualification operationalization [Electronic Version]. *Applied Psychology in Criminal Justice, 13*(1), 18-32.

ENDNOTES

1. In Oregon, Texas, Washington, and Utah, jurors consider "special issues" in lieu of aggravating circumstances, but the basic framework is essentially the same (Palmer, 2014; see *Jurek v. Texas*, 1976)

2. Technically, each of these death qualification standards is applicable to jury selection for any capital punishment case. The trial judge has discretion in interpreting and applying these standards (*Uttecht v. Brown,* 2007); therefore, there is presumably inter-judge and/or inter-jurisdictional differences. Death qualification standards outline what a prospective juror *may* be excluded for, not what a prospective jurors *must* be excluded for (Swafford, 2011). A common critique of death qualification is the ambiguity of death qualification standards that allows for too much discretion (e.g., Swafford, 2011; Thompson, 1989). Although there is no direct evidence specifically of inter-jurisdictional variation in the interpretation and application of death qualification standards, there is indirect evidence. For instance, one study of capital cases found that judges in Kentucky were more likely to sustain defense challenges compared to judges in South Carolina (Nietzel, Dillehay, & Himelein, 1987). When the defense challenges for cause based on death qualification and the judge does not sustain it, this can be cited as grounds for an appeal (e.g., *Chappell v. State*, 2009).

3. *Morgan* and *Witherspoon* are often interpreted as "reverse-*Witt*" standards (see Haney et al., 1994). That is, by indicating they would always vote for a death penalty or never vote for the death penalty, prospective jurors are *ipso facto* indicating their performance as a capital juror would be impaired and are therefore excludable according to the *Witt* standard (McNally, 1985; *Morgan v. Illinois,* 1992).

4. Butler and Moran (2002) did not call it the *Witherspoon-Morgan* standard or explicitly address *Morgan v. Illinois* (1992), but nonetheless incorporated *Morgan* it into their *Witherspoon* standard (i.e., they deemed excludable jurors who would never vote for the death penalty *and* jurors who would always vote for the death penalty).

5. Dillehay and Sandys (1996) stress that, in their experience, jurors are not excluded based on the answer to one question alone. Increased questioning likely exacerbates the biasing effects of death qualification (see Haney, 1984, 2005) and some scholars advocate for a one question death qualification. "A single question in *voir dire*, asking whether prospective jurors have such strong feelings about the death penalty one way or the other that they would not be able to obey their instructions and oath, is sufficient to meet the standard in *Witt.* More is not only unnecessary, but also actively stacks the deck against capital defendants, and should stop" (Rozelle, 2002, p. 694).

6. Applying various combinations of the standards yielded results that tended to mimic the results of the *Witherspoon-Morgan* standard or the *Witt* standard. The participants death qualified according to the *Witherspoon-Morgan* standard *or* the *Witt* standard did not yield significantly different sentencing decisions ($\chi^{(1)} = 2.66$, $\phi = -0.09$, $p = .10$); however, participants death qualified according to the *Witherspoon-Morgan* standard *and* the *Witt* standard were more likely to render a death sentence and less likely to render a life sentence ($\chi^{(1)} = 4.04$, $\phi = -0.09$, $p = .04$). Taken collectively, these findings suggest the *Witherspoon-Morgan* standard is the "driver" of the death qualification effect on sentencing decisions.