BEYOND *ATKINS:* HOW DO THE PRONGS PERFORM DURING SENTENCING?

by

MARY E. WOOD

KAREN L. SALEKIN, COMMITTEE CHAIR
STANLEY L. BRODSKY
FRANCES CONNERS
WILLIAM HART
HONORABLE BRAD ALMOND

A DISSERTATION

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

TUSCALOOSA, ALABAMA

2016

ABSTRACT

In *Atkins v. Virginia* (2002) the High Court categorically excluded individuals with intellectual disability (ID) from a punishment of death due to the limited judgment, poor reasoning, and reduced levels of impulse control inherent in the disability. This research explored the impact of the diagnostic prongs during sentencing for offenders found guilty of a capital crime and who failed to prove ID during an *Atkins* hearing. The current study used a mock jury deliberation paradigm with a large sample of undergraduate students divided into four-, five-, or six-person mock juries. Two of the three diagnostic prongs, limited intelligence and deficits in adaptive behavior, were manipulated with the goal of identifying how these deficits (or lack thereof) are interpreted independently, and in conjunction with, one another during capital mitigation. The results indicated that both IQ and AB deficits are considered mitigating by death-qualified mock jurors, and information about deficits in one or both of these areas was associated with a 1.8 times greater likelihood of a sentence of life without the possibility of parole relative to the condition in which neither IQ nor AB deficits were present. Consistent with the High Court's ruling in *Atkins v. Virginia* (2002), participants who believed the hypothetical defendant had ID were significantly less likely to sentence him to death as opposed to life in prison without the possibility of parole. More broadly speaking, the current study also provided evidence that perceptions of mitigating factors mediate the relationship between individual attitudes and ultimate sentencing determinations, and perceptions of mitigating factors can be understood

through the lens of attribution theory. Implications are discussed with a particular emphasis on how this information can be used in the courtroom. Recommendations for research are offered.

DEDICATION

For my dad, JBS – never forget, grace happens.

LIST OF ABBREVIATIONS AND SYMBOLS

β	Beta: Regression coefficient, the average amount by which the dependent variable increases with unit increases in the independent variable; the slope of a line
df	Degrees of Freedom: Number of values in a final calculation that are free to vary (i.e., number of independent observations minus the number of estimated population parameters)
F	*F* statistic: Value calculated by the ratio of two sample variances
M	Mean: The sum of a set of values divided by the number of values in the set
n	Sample size of a group
η_p^2	Partial Eta-Squared: Value of effect size, the proportion of the total sum of squares that is explained by the predictor(s)
OR	Odds Ratio: Relative odds of an occurrence given a particular condition of interest
p	Probability associated with the occurrence under the null hypothesis of a value as extreme as or more extreme than the observed value
phi	Measure of effect size for chi-square statistic
R^2	R-squared: Coefficient of determination, a measure of effect size
SD	Standard Deviation: Value of variation from the mean
SE	Standard Error: Measure of the statistical accuracy of an estimate
t	*T* statistic: Value determining whether sample means differ
χ^2	Chi-square test of significance of model fit
z	*z* statistic: Measurement of a score's relationship to the mean in a group of scores in standard deviation units
<	Less than
=	Equal to

+/-	Plus or minus
%	Percentage

ACKNOWLEDGEMENTS

There were many people who were instrumental in the completion of this project. The members of my dissertation committee, Drs. Karen Salekin, Stan Brodsky, Will Hart, Fran Conners, and the Honorable Brad Almond, provided important and useful feedback during the initial stages of this project, and supported my continued consultation with them over its course. I have also had a number of incredible faculty mentors and supervisors over the years who have inspired and encouraged me to work hard and dream big, especially Drs. Clay Shealy, Jeff Helms, and Amber Simpler. Several of these supervisors also encouraged my development as a clinical psychologist, further facilitating my interest in and passion for clinical work. Finally, and perhaps most importantly, this project would not have been possible without the unending love and support from my closest family and friends. Without them and their countless demonstrations of support, this project would never have been completed. My husband, the almost astronaut, deserves a special shout out for all of his hard work in getting me here – this accomplishment is almost as much his as it is my own.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Death Penalty in the United States

Since the first execution took place in 1608 in Virginia, the death penalty has endured a long and sordid history in the United States. Early on, the death penalty was an available punishment for minor crimes (e.g., stealing grapes) and executions were generally conducted in public. Throughout history, however, the number of crimes for which death was an available punishment decreased, mandatory sentences were abolished, and increasingly humane (and more private) methods of execution have been introduced. Despite these improvements, the death penalty continues to be challenged and issues remain with respect to its constitutionality. Most notably, in 1972 (*Furman v. Georgia*), the Supreme Court of the United States (SCOTUS, or the High Court) deemed the death penalty unconstitutional and did so based on their view that it was a violation of a defendant's constitutional rights as allowed by the 8th amendment (i.e., protection from cruel and unusual punishment). The High Court relied on evidence of arbitrary sentencing practices in capital cases that were seemingly based solely on a defendant's race. Writing for the majority, Justice Stewart commented that capital punishment should not be awarded on the "constitutionally impermissible basis of race," of which there was evidence for at the time, and further noted it was intolerable that capital punishment could be and was "so wantonly and so freakishly imposed."

The death penalty was reinstated in the United States just four years later following the collective *Gregg* decisions (i.e., *Gregg v. Georgia,* 1976; *Jurek v. Texas,* 1976; & *Proffitt v.*

Florida, 1976) in which reforms were introduced including the institution of separate guilt and sentencing phases of capital trials (i.e., a bifurcated process) and an automatic appeals process for those convicted of a capital crime. The reform most relevant to the current research, however, was the proposal of guided discretion with respect to sentencing (i.e., the fact-finder would be "provided with standards to guide its use of the information" presented during sentencing; *Gregg v. Georgia*, 1976). This move toward increasing discretion of the fact-finder opened the door for individualized consideration of aggravating and mitigating factors during sentencing.

Initially proposed in the *Gregg* decision, the court decided that "the jury is not required to find any mitigating circumstance in order to make a recommendation of mercy… but it must find a statutory aggravating circumstance before recommending a sentence of death" (1976). To sum, the court ruled that capital punishment should only be available in the presence of an aggravating circumstance, and that the fact-finder must consider opposing mitigating circumstances in making sentencing determinations. In other words, fact-finders in capital cases must consider the evidence for (aggravation) and against (mitigation) a sentence of death for a defendant; the former is a requirement for a sentence of death to be imposed, while the latter is not. Two years later, SCOTUS extended the *Gregg* decision and expanded the domain of mitigating factors to include "any aspect of a defendant's character or record and any of the circumstances of the offense that the defendant proffers as a basis for a sentence less than death" (*Lockett v. Ohio*, 1978).

Since the reinstatement of the death penalty in 1976, certain classes of offenders have been categorically excluded from imposition of the death penalty, which Cunningham (2005) has argued is a demonstration of the High Court's increasing emphasis on individual characteristics of capital defendants. In 1988, the High Court excluded those under the age of 16 at the time of

the offense (*Thompson v. Oklahoma*, 1988) and later extended protections for those up to age 18 (*Roper v. Simmons*, 2005). Most relevant to the current research, in 2002 SCOTUS determined that the execution of individuals diagnosed with intellectual disability (ID; formerly mental retardation or MR) was unconstitutional due to deficits in reasoning, judgment, and impulse control and the ramifications thereof (*Atkins v. Virginia*, 2002).

Atkins

Although SCOTUS referenced the definitions of ID provided by the American Psychiatric Association (APA) and the American Association on Intellectual and Developmental Disabilities (AAIDD) in the *Atkins* decision, the method and criteria for determining whether a defendant met criteria for ID was left up to the states. Both definitions follow a three-pronged diagnostic classification system which includes deficits in intelligence (IQ), deficits in adaptive behavior, and an onset of the first two prongs during the developmental period. Although definitions of both agencies have been modified since *Atkins*, the three-pronged system remains in place and both currently align in terms of their conceptualization of deficits and age of onset (APA, 2013; AAIDD, 2010). As of 2007, all but four of the states that permitted capital punishment specifically mentioned two of the three diagnostic prongs required for a diagnosis of ID in *Atkins* statutes; the remaining four states noted "presumptive evidence" of adaptive behavior deficits based on a particular IQ score (DeMatteo, Marczyk, & Pich, 2007). Although most states mentioned two of the three prongs, few states provided sufficiently descriptive criteria by which to define ID (i.e., poor operational definitions of the criteria), and less than half of death penalty states provided a clinical definition of ID that mapped onto widely accepted diagnostic criteria (DeMatteo et al., 2007). Perhaps most concerning was the definition provided

by Kansas, whereby the “cognitive and volitional elements” inherent in insanity statutes were included as definitive of ID.

Although ID is the only psychiatric diagnosis which has resulted in categorical exemption from punishment, limits to its successful implementation still exist. As mentioned before, states are not uniform in their application of *Atkins* (i.e., only 11 states defined ID using well established and accepted clinical criteria; DeMatteo et al., 2007). Further, it was not until 2014 that SCOTUS ruled that states had to consider the standard error of measurement (SEM) in interpreting IQ score estimates (*Hall v. Florida,* 2014). Since *Atkins,* at least 371 individuals have brought claims before the Court requesting review under *Atkins.* The percentage of success in these cases is difficult to know, but based on data from Blume, Johnson, Marcus, and Paavola (2014), *Atkins* claims are successful in the range of 44% to 55% of the time, and generally exceed estimated rates of success for other capital claims (e.g., competence for execution). Of the 371 filed claims between 2002 and 2013, approximately half failed on all three diagnostic prongs, approximately one-third failed to prove deficits in intellectual functioning, 12% failed to prove deficits in adaptive behavior, and about 2% failed to prove that the observed deficits were present during the developmental period (Blume et al., 2014). In their discussion of adaptive behavior, Blume and colleagues (2014) noted that “stereotypes and general misunderstandings about what people with intellectual disabilities can achieve are likely the most significant factors affecting Prong 2 losses" (i.e., failure to prove adaptive behavior deficits; pg. 408). Further, “*all* of the reported losses on Prong 2 from 2008 to 2012 discussed some aspect of the claimant’s prison behavior as support for the court’s conclusion that the claimant failed to demonstrate deficits in adaptive functioning sufficient to satisfy Prong 2” (Blume et al., pg. 406).

Blume and colleagues (2014) also found "jurisdictional variations" in the success of filed *Atkins* claims, with the rate of success in Alabama falling at 12% and success in North Carolina reaching 80%. The discrepancy among states was attributed to the availability of funding allotted for capital cases in addition to the degree of deviation from accepted clinical standards (i.e., statutes that more closely matched accepted definitions corresponded with higher rates of success). There was also evidence that rates of success varied upon the fact-finder as juries tended to be more stringent in their application of *Atkins* relative to judges; Blume et al. (2014) postulated that judges' increased ability to set aside heinousness and also timing of the claim (i.e., pre- versus post-trial) may explain this difference. In short, a review of historical *Atkins* cases suggests some inconsistency in terms of determinations of ID that may be the result of extra-clinical factors (e.g., stereotypes and misunderstandings of ID; consideration of irrelevant information), statutory differences, and/or bias of the fact-finder.

Capital defendants who fail to prove ID under *Atkins* are not awarded protections and are treated similarly to any other capital defendant. Of concern, the difference between having ID and not having ID may simply be a few IQ points or deficits in adaptive functioning that were not deemed to be significant enough, either by test scores and/or information gleaned from third party sources. In a sense, the diagnostic prongs for establishing ID are arbitrary cut-offs that rely upon statistical abnormality (e.g., an IQ score more than two standard deviations below the mean, or the bottom 2.3% of the population) and interpretation of the trier-of-fact. This issue of arbitrary cut scores is further complicated by the fact that IQ scores often differ depending on the test used (Simon & Clopton, 1984; Spitz, 1986; Silverman, Miezejeski, Ryan, Zigman, Krinsky-McHale, & Ury, 2010) and tend to fluctuate over time as a result of error and/or practice effects (Larsen, Hartman, & Nyborg, 2008; Whitaker, 2008). As such, the difference between

individuals' IQ scores may reflect a variation in testing or temporal fluctuations which may lead to ineligibility for *Atkins* relief. This criticism was further articulated by Greenspan and Switzky (2003) who argued that exemption from capital punishment should be determined on the basis of demonstrated social vulnerabilities and not the diagnostic label, as the vulnerabilities themselves (not the label) were the basis of the original *Atkins* exemption (i.e., "is it not both more efficient and equitable to exempt from execution those who have those vulnerabilities, irrespective of whether they are eligible for the label of MR?," Greenspan & Switzky, 2003, pg. 20).

For the defendants who fail to prove ID in an *Atkins* hearing, there likely exists some support for ID. For relief under *Atkins*, the IQ score(s) almost invariably need to be at or below 75, leaving very little room for error. There is some evidence that the Courts are willing to accept scores above the cut-off in light of information on the Flynn Effect, practice effects, and SEM as 46% of successful *Atkins* claims through 2013 (n = 49) had at least one IQ score over 75 and one-fifth had at least one IQ score over 80 (Blume et al., 2014). The current research is designed to investigate how two of the three prongs for ID, deficits in IQ and AB, are evaluated by jurors during mitigation following a failed *Atkins* claim, as it is unknown how these factors are perceived by fact-finders. Identifying how these "near misses" are interpreted by jurors is relevant to defense and prosecuting attorneys who may be in a position to present and/or refute evidence of mitigating factors presented during sentencing.

Capital Sentencing

For individuals found criminally responsible (i.e., guilty) of a capital offense, the job of the defense is to pursue evidence that the defendant is less morally culpable for the offense than other capital defendants and, therefore, deserving of a sentence of life imprisonment without the possibility of parole (LWOP) rather than death. This process, referred to as capital mitigation,

charges the fact-finder (i.e., either the judge or a jury) with the task of weighing aggravating and mitigating factors to determine the appropriate sentence for the defendant. When discussing mitigating factors, Schroeder, Guin, Pogue, & Bordelon (2006) noted the following:

> …the gold standard for mitigation practice is reliance on focused investigation and presentation of evidence about a defendant's life and character that is sufficient to compel jurors to consider a complexity of personal, social, and ecological issues in deciding attribution of responsibility that mitigate defendants' legal accountability in capital offenses (pg. 355).

As established in *Lockett v. Ohio* (1978), the domain of mitigating factors (i.e., facts that indicate a sentence of LWOP) is virtually limitless. In *Lockett,* the High Court specifically stated the following:

> We conclude that the Eighth and Fourteenth Amendments require that the sentencer, in all but the rarest kind of capital case, not be precluded from considering, as a mitigating factor, any aspect of a defendant's character or record and any of the circumstances of the offense that the defendant proffers as a basis for a sentence less than death.

In addition to the limitless nature of mitigating factors, states provide factors which may be considered by the fact-finder during the sentencing phase. Alabama includes the following examples of mitigating factors in the state's code: (1) no prior criminal history, (3) the victim was a participant in the conduct, and/or (7) the age of the defendant at the time of the crime (Ala. Stat. Ann. 13A-5-51 and 13A-5-52); in all, Alabama lists seven statutory mitigating factors.

In contrast to mitigating factors, aggravating factors are those that support a sentence of death. In capital trials, the prosecution must prove at least one aggravating factor beyond a reasonable doubt for death to be an available punishment (i.e., a factor that reasons for a sentence

of death; *Ring v. Arizona*, 2002). Examples of aggravating factors include such things as the heinousness of the offense, the commission of the offense during a felonious act, and/or the prior criminal history of the defendant (Ala. Code 1975, Section 13A-5-45). In all, Alabama lists ten statutory aggravating factors for the fact-finder(s) to consider.

In a systematic study of qualitative data from interviews with capital jurors across the country, 97.5% of the over 650 capital jurors surveyed reported that they remembered sentencing deliberations "fairly well" or "very well," while 92.5% remembered mitigation testimony to the same degree (Bowers, 1995). Guilt-phase participation strongly influenced sentencing decisions for these jurors, as just under half (49.8%) were undecided at the start of the sentencing phase. Put differently, half of capital jurors made determinations of sentence prior to entering the sentencing phase of the capital trial (Bowers, 1995). This finding may be one contributing factor to the belief that "nothing works" when it comes to mitigation (Eisenberg, 2004).

Haney, Sontag, and Costanzo (1994) succinctly described the veritable difference between capital mitigation and typical criminal responsibility determinations as more "infused with values" and as "uniquely a process of *moral assessment*" (pg. 153). Further, Haney and colleagues (1994) commented on the "unbounded" and "enormous and unprecedented" nature of testimony, absent a structural framework from which to base decisions, as the primary characteristics which distinguish penalty phase testimony from determinations of guilt and/or criminal responsibility. Consistent with their assessment, it is important to investigate the process by which jurors arrive at ultimate decisions given the "unbounded" amount of information and the relative absence of a framework to guide decision-making. In essence, capital mitigation is an unguided, subjective, and oft misunderstood process where grave matters of life and death are decided upon; it is the gravity of these determinations and the fact that above all else, death is

"qualitatively different" (*Lockett v. Ohio*, 1978; *Wainwright v. Witt*, 1985; *Atkins v. Virginia*, 2002), that makes research in this area an imperative endeavor.

Research on Mitigating Factors

Previous research has investigated the process by which fact-finders interpret mitigating evidence during capital sentencing and also the mitigating impact of certain factors. With regard to specific mitigating factors, Garvey (1998, as cited in Sandys et al., 2009) divided mitigating factors into four categories that ranged from residual doubt (the most powerful) to relative culpability (the least powerful of the four categories). Garvey (1998) identified personal factors inherent to the individual (i.e., uncontrollable) that are related to perceptions of decreased culpability, including ID and mental illness, as falling the proximate culpability category. These factors were the second most effective grouping of factors based upon ratings of capital jurors out of South Carolina. The most effective yet least prevalent category was residual doubt (i.e., doubts about the defendant's guilt). As one might expect, residual doubt should be rare once the process of capital mitigation has been set in motion, as a finding of guilt has already been decided upon and accepted by the court.

Previous research with undergraduate and graduate students revealed that mitigating evidence is effective in the sense that death sentences were less likely in cases in which certain mitigating factors were presented than in cases where these factors were not presented (Barnett et al., 2004). Participants were presented with ten vignettes with and without mitigating evidence. Mitigating factors included serious mental illness (schizophrenia), drug use, borderline mental retardation, and a history of abuse or migraine headaches, among others. Results indicated that information about serious mental illness and borderline mental retardation were associated with an increased likelihood of a life sentence rather than death (Barnett et al., 2004). In addition,

information about abuse and drug use yielded a higher proportion of life sentences than the vignettes in which “no intentional psychological/psychosocial mitigating evidence” (Barnett et al., 2004, pg. 757) was presented, though the same factors were rated as aggravating when presented independent of a vignette. The authors posited that context may have influenced ratings and/or that information about drug and alcohol use is considered to be aggravating; the latter has been supported by other research presented elsewhere (Barnett et al., 2004).

In addition to the content of mitigating factors, Tetterton and Brodsky (2007) found that the sheer number of factors influences perceptions in mitigation. Undergraduate students were presented with a fictional case coupled with various iterations of one, two, or three mitigating factors. Results showed the impact of mitigating factors was higher when three mitigating factors were presented than when only one or two mitigating factors were presented. The researchers also found that lengthier periods of physical abuse were perceived as more mitigating than less lengthy periods (i.e., one years versus five or ten years; Tetterton & Brodsky, 2007). The authors went on to note that experts should be inclusive and comprehensive when testifying to these issues, as details such as length of abuse and the number of risk factors can influence perceptions.

In a more focused investigation of mitigating factors, Barnett et al. (2007) presented ten mitigating factors to a sample of undergraduate students and a sample of community residents. Each participant was asked if each individually presented factor would result in a sentence that was more lenient, harsh, the same, or unknown relative to if the factor was absent. Factors which resulted in more lenient sentences included ID, a history of psychiatric hospitalizations or schizophrenia, head injuries, and no prior criminal record (Barnett et al., 2007). In general, information about childhood physical and sexual abuse did not alter sentences, while information

about drug and alcohol intoxication and addiction was perceived as aggravating by 52% and 38% of the sample, respectively. Subsequent research has replicated the latter finding (Bjerregaard, Smith, Fogel, & Palacios, 2010; Lynch & Haney, 2009), and will be discussed subsequently as the backfire effect.

In an effort to identify influential factors that might predict juror decision-making in real world settings, research has also analyzed the content of actual capital cases (e.g., Wolbransky, 2011; Bjerregaard et al., 2010; Gillespie, Smith, Bjerregaard, and Fogel, 2014). One downside to this type of research is the lack of experimental control, but the advantage is the ability to identify factors which appear to be influential in real capital cases. This research has generally focused on capital mitigation in North Carolina whereby only one capital juror must agree that a factor is mitigating in order for it to be considered as a mitigating circumstance during sentencing and therefore marked "yes" on a required capital jury form (i.e., *McKoy v. North Carolina*, 1990). Prior to *McKoy* (1990), mitigating factors required unanimous agreement to be considered by the capital jury during sentencing. Previous research, presented below, has benefited from this requirement given the opportunity to analyze responses on the "Issues and Recommendations" form completed by capital juries in North Carolina indicating which factors were accepted versus rejected.

Wolbransky (2011) was interested in the influence of mental illness-related mitigating factors, operationally defined as (1) mental or emotional disturbance, (2) impaired appreciation of criminality or ability to conform conduct, or (3) any evidence of a mental illness.[1] At least one

[1] The first two are statutory mitigating factors which appear in North Carolina's criminal procedure. The first refers to the "influence of mental or emotional disturbance" in the commission of the crime, while the second refers to "the capacity of the defendant to appreciate the criminality of his conduct or to conform his conduct to the requirements of the law." The third, non-statutory factor refers to evidence that was presented related to a mental illness undefined by the first two, statutory factors.

of these mental illness-related mitigating factors was presented in over three-quarters of the 326 capital cases in North Carolina between 1982 and 1998, and was successful approximately 73% of the time. The most frequently occurring dyad was the presentation of (1) mental or emotional disturbance and (2) inability to conform one's conduct (30.7% of cases). Evidence of the first statutory factor, "mental or emotional disturbance," was the most successful of the mental illness-related factors as it was accepted just over 70% of the time it was presented. The non-statutory factor, evidence of a mental illness, was presented in over one-quarter (27.6%) of the cases and accepted almost 60% of the time. Similarly, evidence of a severe mental illness (e.g., a psychotic spectrum illness), was submitted in about 7% of cases and accepted approximately 65% of the time. When the first factor, emotional or mental disturbance, was rejected by the jury, there was an increased probability of a death sentence relative to when it was either accepted or not presented at all. Further, "a death sentence was about twice as likely when the jury agreed with "mental or emotional disturbance" as a mitigating factor than when this factor was not presented to the jury" (Wolbransky, 2011, pg. 20).

Acceptance of the other statutory mitigating factor, ability to conform/appreciate conduct/criminality, was associated with a lower probability of a death sentence than when it was rejected or not submitted. These findings suggest that capital jurors discriminated between a mental illness and a mental illness that explicitly influenced the defendant's thinking or judgment, potentially providing a link between the mental illness and crime, rather than having to make that link themselves from only the presence of a mental illness. In other words, when the capital jury was provided with explicit information about the influence of one's mental state on their ability to conform their conduct (i.e., an inability to conform conduct), death sentences were less likely, whereas evidence of emotional disturbance presented alone was associated with an

increased likelihood of death. Again, this finding is consistent with the idea of the backfire effect whereby factors intended to be mitigating have been used to support a sentence of death.

In a review of 834 capital cases out of North Carolina from 1990 through 2009, Gillespie et al. (2014) similarly found that mental health mitigation (operationally defined as mental disturbance, capacity to appreciate criminality, low IQ/diminished capacity, and/or a specific mental disorder), was accepted at least half of the time it was submitted. Results of a logistic regression model suggested that acceptance and rejection of these factors explained approximately half of the variability in the ultimate sentences awarded for capital defendants, even after controlling for a number of relevant variables. When the capital jury rejected evidence of a defendant's specific mental illness as a mitigating factor, there was a one-third increase in the likelihood of receiving a sentence of death. However, acceptance of this factor was not associated with ultimate sentence. Gillespie and colleagues (2014) concluded that this finding supported the idea of 'converted mitigation' (originally coined by Haney et al., 1994, discussed subsequently) whereby the jurors used evidence that was presented as mitigating to support a sentence of death, therefore converting it to an aggravating factor. They went on to note that one possible reason for this conversion was that of stereotyped ideas about the relationship between mental illness and dangerousness, therefore using evidence of a mental illness to support their decision to vote for death.

As seen from the above studies, factors which decrease the culpability of the defendant and are outside of the individual's control (e.g., mental illness, ID, history of abuse), are mostly perceived as mitigating, while factors within the control of the defendant (e.g., drug and alcohol use) are generally perceived as aggravating. It follows then, that information about deficits related to the first two diagnostic prongs of ID – deficits in cognitive ability and adaptive

functioning – would be considered mitigating, as these factors would be attributable to a disability that is outside of the individual's control. However, when presented without an attributable cause (i.e., a diagnosis of ID), individuals may perceive these factors more variably and perhaps within the control of the defendant, and therefore more aggravating. The current research has been designed to investigate how these factors are interpreted during mitigation when they are presented without with a diagnostic label, ID. Presenting these factors independent of a diagnosis and focused only on the resulting vulnerabilities is consistent with the argument of Greenspan and Switzky (2003) that the diagnostic label is arbitrary and less important than demonstrated deficits when making determinations of culpability. At present, however, it is unknown how these vulnerabilities are interpreted by fact-finders in a capital context – the focus of the present research.

Mitigation defenses. One impediment to capital mitigation receptivity is that of the belief that capital mitigation offers only an excuse, rather than an explanation, of behavior (Eisenberg, 2004). This belief has led some individuals, lawyers included, to report that "nothing works" when it comes to capital mitigation. Because of this, the way in which information is presented to the fact-finder(s) is important with respect to if and how that information is received. As such, previous research has investigated capital mitigation defenses (i.e., the way in which information is presented). Potential defense strategies include reliance upon a story model,[2] or walking the fact-finder through how such an individual became someone who could be charged with capital murder. Within the confines of the story model, focus can be placed on particular defense strategies such as appealing to the belief that the death penalty is outdated and

[2] The story model involves a "comprehensive and empirically well-documented understanding of the capital defendant's life" (Haney, 2008). This story should be chronological and work through the elements of the defendant's life to create a well-crafted story of their life.

immoral, or that the defendant was so severely incapacitated via mental illness or substance use that a sentence of death is inappropriate (White, 1987; Trahan, 2011).

Previous research has indicated that mock jurors were more receptive to anti-capital punishment sentiments than a defense about mental illness (i.e., that the defendant was severely mentally ill; White, 1987) as evidenced by more punitive decisions in the mental illness conditions. Although punitive judgments were highest in response to the mental illness defense, ratings of the volition of the defendant were lowest in this condition (White, 1987). Of particular concern was that "some mock jurors reacted very strongly against a "mental illness" explanation of the defendant's behavior" (White, 1987, pg. 125) as evidenced by statements such as "mental illness is no excuse" and that the "defendant did not seek help for his problems." Mental illness as a defense is not the only content of mitigation which has been shown to be interpreted negatively by mock jurors as similar findings have been published related to alcohol and substance use mitigation (see subsequent section on the backfire effect; Barnett et al., 2007; Bjerregaard, et al., 2010; Lynch & Haney, 2009).

In a thematic analysis of interviews with actual capital jurors, Trahan (2011) found that of 916 capital jurors, only 76 "discussed their impressions of mitigation at least once during their interview" (pg. 3). Five themes mentioned by jurors included (1) questioning the guilt of the defendant, (2) requests for mercy, (3) claims of diminished capacity secondary to mental illness, mental retardation or substance use, (4) background of the defendant, and (5) a multifaceted approach in which multiple strategies (i.e., themes) were used (Trahan, 2011). Interviews with capital jurors were coded in order to identify relevant reactions to these themes. Similar to the findings of White (1987), "diminished capacity – was rejected by over 90% of the jurors" (Trahan, 2011, pg. 11). The author suggested that education regarding mental illness and mental

retardation (or ID), particularly the link between the disability and one's behavior, may minimize the negative reactions to this defense strategy (i.e., mental illness and ID testimony "should be detailed, specific, informative, and substantiated by expert testimony and/or evidence," Trahan, 2011, pg. 12). The jurors were most receptive to a multifaceted approach which combined multiple themes, suggesting that reliance upon several approaches (e.g., presenting about the defendant's background along with explicit and detailed connections between mental illness/retardation and its relevance) will be the most effective during mitigation.

In addition to the content of the testimony, a review of research from the Capital Jury Project revealed negative reactions to professional experts on the stand, and found that more positive reactions resulted when experts integrated case-specific information into their testimony (Devine et al., 2001). Similarly, Keene (2010) recommended that experts incorporate case-specific facts into their testimony in order to increase the likelihood of being viewed as credible. In short, capital mitigation experts should (1) rely on several thematic approaches to their testimony and (2) tie the content of that testimony directly to the case in which they are testifying; doing so increases juror receptivity to information and increases the likelihood they (the expert) will be viewed as credible.

Intellectual disability, the diagnostic label. Previous research has found that information about ID (i.e., the diagnostic label) is considered mitigating in the minds of mock jurors (Barnett et al., 2004; Barnett et al., 2007). Over 70% of participants surveyed said that they would vote for a more lenient sentence in the presence of a diagnosis of mental retardation (the appropriate diagnostic label at that time; Barnett et al., 2007). In addition to being perceived as mitigating, a juvenile defendant with mild ID was perceived as less deviant and less responsible for a crime than an identical defendant of average intellect (Najdowski, Bottoms, &

Vargas, 2009). In a second, related study, Najdowski et al. (2009) found that seriousness of the offense interacted with disability status such that ID influenced ratings of guilt and deviance for less serious crimes. This interaction has been demonstrated elsewhere (Gibbons, Gibbons, & Kassin, 1981; Bottoms, Nysse-Carris, Harris, & Tyda, 2003) and researchers have theorized that the patronization effect (the assignment of external attributions to individuals with ID) triggers the belief that disabled individuals are less sophisticated and therefore less likely to commit more serious offenses. However, Najdowski et al. (2009) noted that in the absence of external attributions for a crime, mock jurors may "fail to consider" or "ignore" an offender's disability in arriving at a sentence. In other words, because individuals with ID are perceived as unlikely to commit serious offenses (or perhaps unable to do so), individuals may look to external factors such as the environment or history of abuse.

Although the mitigating influence of the diagnostic label of ID has been documented in mock jury studies (Barnett et al., 2004; Barnett et al., 2007; Najdowski et al., 2009), it is currently unknown how the specific impairments associated with ID (i.e., deficits in IQ and AB) are evaluated independently and in conjunction with one another absent a diagnostic label. Because research has indicated that perceptions of ID reflect stereotyped connotations of the diagnostic label itself rather than knowledge of impairments which comprise the disability (see Scior, 2011, for a review), the mitigating value of these factors independent of a diagnosis may be unrelated to perceptions of the label alone. In other words, it may be that the diagnostic label "ID" is mitigating secondary to inaccurate perceptions of the disability instead of reflecting the mitigating value of the behaviors which characterize the disorder (i.e., deficits in IQ and AB).

Low IQ. IQ, broadly speaking, is a measure of problem-solving ability and reasoning skills (APA, 2013). As previously mentioned, in the case of *Atkins v. Virginia* (2002), the High

Court emphasized the importance of intellectual deficits which make an offender less culpable for a capital offense. In the majority opinion, Justice Stevens wrote the following: "mentally retarded persons frequently know the difference between right and wrong and are competent to stand trial, but, by definition, they have diminished capacities to understand and process information, to communicate, to abstract from mistakes and learn from experience, to engage in logical reasoning, to control impulses, and to understand others' reactions." In *Atkins,* the Court categorized individuals with ID as naïve defendants, confessors, and offenders, therefore at risk of exploitation and harm at all stages of the trial process (Greenspan & Switzky, 2003). In addition to the High Court citing the importance of IQ, research has indicated that IQ is linked to criminality in general and that the IQ scores of incarcerated populations are lower than populations who are not incarcerated (Kane, 2003; Dwyer & Frierson, 2006), suggesting that IQ is an important factor for individuals involved in the criminal justice system. In fact, low IQ is a known risk factor for juvenile delinquency that has been documented in the literature (U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention, 2004; Murray & Farrington, 2010).

To date, very little research has been conducted on the influence of IQ during capital mitigation. Gillespie and colleagues (2014) recently found that low IQ was submitted as a mitigating factor in 22.9% of capital cases in North Carolina between 1990 and 2009, and was accepted in just over half of those cases. As a reminder, in North Carolina mitigating factors need to be considered by only one capital juror to be accepted. In cases where low IQ was accepted, the proportion of death sentences was slightly lower relative to cases whereby it was rejected (i.e., 45.5% versus 65.2%). Neither acceptance nor rejection of low IQ as a mitigating factor was associated with recommendations for death, however. The authors attributed this lack

of an association to confusion related to the difference between ID and mental illness (see Sandys et al., 2008), as well as a lack of consensus with respect to which defendants have ID.

Adaptive behavior deficits. Adaptive behavior, broadly defined, is the collection of behaviors necessary to maintain functional independence based upon culturally expected norms (AAIDD, 2010; APA 2013). Adaptive behavior was first included in the diagnostic criteria for ID in 1877 and "referred to by one author as "moral monstrosities"" (Kerlin, 1877, pg. 37, as cited in Bruininks, Thurlow, & Gilman, 1987, pg. 70). The term "adaptive behavior" arose in the 1960s, but it was not until 1973 that "adaptive behavior became a requirement rather than just a consideration in the diagnosis of mental retardation, even though IQ scores continued to have the greatest weight" (Stevens & Price, 2006, pg. 3). Since its inclusion, the definition of adaptive behavior has been modified several times with the most recent conceptualization dating to 2010 when the AAIDD changed from conceptualizing adaptive behavior as falling into ten skill areas to three domains. At present, adaptive behavior refers to how the individual handles situations in both a personal and social context, as compared to generally agreed upon standards of behavior (AAIDD, 2010; APA, 2013). Significant deviations from typical behavior expected for the specified age and culture of an individual may reflect deficits in adaptive behavior.

With regard to the influence of adaptive behavior information in the courtroom, in a review of 19 pre-*Atkins* transcripts in Texas, Kan, Boccaccini, McGorty, Noland, and Lawson (2009) found that adaptive behavior (i.e., information about the defendant's social history, planning abilities, academic performance, etc.) was referenced more often than IQ or the diagnostic label at that time (MR) in general. In fact, almost three-quarters of information referenced regarding ID was about adaptive behavior. Every case addressed at least one of the three domains, and 14 of the transcripts addressed all three areas; not surprisingly, the adaptive

behavior areas addressed were most often those consistent with lay perceptions of ID (e.g., functional academics and self-direction). Even further, of the collection of cases that failed to prove ID during *Atkins*, one-third utilized prison behavior as evidence of adaptive behavior despite the artificial setting (Blume et al., 2009a. & 2009b.), and over half of *Atkins* transcripts mentioned criminal behavior as evidence of adaptive behavior despite the perspective of its inappropriateness (Kan et al., 2009; Tassé, 2009). Further, all of the failed *Atkins* claims between 2008 and 2013 mentioned prison behavior and generally viewed prison employment and "normal" prison behavior as evidence of adaptive functioning (Blume et al., 2014). As can be seen from the above, there are unresolved issues in the conceptualization, assessment, and interpretation of adaptive behavior in the forensic context. At present, it is unknown how adaptive behavior information may be interpreted by a fact-finder during the course of capital mitigation.

The backfire effect. There is research to suggest that factors that have been found to be mitigating in some cases are viewed as aggravating in others (Barnett et al., 2007; Bjerregaard, et al., 2010; Lynch & Haney, 2009). The process whereby these factors are viewed as aggravating despite the intention for mitigation has been called converted mitigation (Haney et al., 1994), and was later termed the *backfire effect* (Barnett et al., 2007). Factors that have been subjected to such a process include a history of alcohol and drug use and addiction (Barnett et al., 2007; Bjerregaard et al., 2010).

In a review of 804 capital sentencing cases in North Carolina (i.e., 1990 through 2006), researchers found that alcohol and drug mitigating evidence was submitted independently and in conjunction with one another in 10.3%, 13.2%, and 20.3% of cases, respectively (Bjerregaard et al., 2010). Alcohol, drug, and dual mitigating factors were accepted in 4.2%, 3.5%, and 13.1% of

cases where they were submitted. Rejection of the factors was associated with an increased probability of a death sentence, as was the acceptance of a drug mitigating factor (Bjerregaard et al., 2010). The latter finding suggests that although the drug factor was accepted as a mitigating factor, it may have been converted to an aggravating factor as it was associated with an increased probability of a death sentence. The authors noted that despite the significant finding, the effect was modest (i.e., it explained less than 10% of the variability), suggesting that other factors were at play.

In a study of capital mitigation deliberations, Lynch and Haney (2009) found evidence for the backfire effect as well. In their study, 14 to 30% of mock jurors converted mitigating evidence to aggravating evidence. Specifically, 30% of jurors in favor of the death penalty interpreted substance intoxication as aggravating, and 23% used information about a history of severe child abuse as supportive of a sentence of death. During interviews with capital jurors, Garvey (1998) found that alcohol and drug use leading to a lack of control over one's actions was perceived as mitigating by less than one-fifth of the sample. Based on these findings, it appears that some factors, like alcohol/drug use and sometimes abuse history, are perceived as aggravating rather than mitigating.

Attribution theory. Attribution theory provides a framework for how we evaluate the causes of others' behavior. In turn, these attributions influence how we respond to such an individual and/or perceive their actions. This is particularly relevant in a punishment context whereby one's attributions may influence punishment decisions. There are three primary attribution dimensions that have been studied in the literature: controllability, stability, and locus of control (Weiner, 1985). Locus of control is directly related to whether one interprets the cause of an event as internal or external to the individual. In other words, were the actions of the actor

responsible for the event (internal) or was the environment and the subsequent reinforcement of others (external) responsible for the event? Similarly, controllability is related to whether the behavior is within the purview or control of the individual. Stability refers to the degree to which the factor or behavior is constant over time.

Researchers have suggested that attributions can help to explain the backfire effect and account for the different ways in which jurors interpret data (Stevenson, Bottoms, & Diamond, 2010; Graham, Weiner, & Zucker, 1997; Shaver, 1985). As noted by Stevenson and colleagues (2010), "the types of attributions people make about the cause of a crime predict their perceptions of a defendant's degree of responsibility and likelihood for future offending" (pg. 3). Internal, controllable, and stable attributions are associated with more punitive judgments than external, uncontrollable, and unstable attributions (Weiner, 2006; Graham et al., 1997; Shaver, 1985).

Relatedly, attitudes toward individuals with ID have been shown to be consistent with attribution theory as well (Panek & Jungers, 2008). In particular, the stated cause of the disability (i.e., genetic versus self- or other-inflicted) was related to evaluations such that attitudes were more positive when the disability was described as of genetic causes (i.e., uncontrollable) than when described as self-inflicted (i.e., controllable). Conversely, attributions regarding the cause of mental illness (i.e., biological versus non-biological) did not influence self-reported desire for social distance in a sample of Midwestern university students (Sears, Pomerantz, Segrist, & Rose, 2011). In other words, beliefs regarding the controllability and cause of mental illness did not influence self-reported acceptability of said individuals, though similar attributions of the causes of ID did influence attitudes.

But what happens in decision-making when a factor is an unexpected combination of these dimensions? For example, IQ is uncontrollable yet internal and stable. In addition, what if a potentially mitigating factor (e.g., AB deficits) is presented absent an internal attribution? It is possible that AB deficits will be considered controllable (i.e., within the purview of the defendant) in the lack of an adequate explanation like intellectual impairment. Though this has not yet been addressed in the literature, a Texas judge attributed a defendant's AB deficits (in the presence of an IQ of 70) to Antisocial Personality Disorder and described these deficits as "attention-getting" rather than indicative of ID (*Williams v. Quarterman*, 2008), suggesting that AB deficits, when attributed to a cause other than ID, may be perceived as an aggravating factor – even in the presence of a low IQ.

Previous research has also investigated the relationship between general support for the death penalty and attributional style. In a focused, vignette-based study with a large sample of jurors in Florida, Cochran, Boots, and Heide (2003) found that attribution style was related to punitiveness and sentencing recommendations such that dispositional attributions for crimes were more punitive and more strongly associated with a sentence of death, while situational attributions were associated with the opposite "for all but mentally retarded and mentally ill offenders" (pg. 80). Demographic variables that that were consistently predictive of support included race, gender, political ideology, and region (i.e., south versus not), though the variability explained by these variables was small (i.e., 9-16%; pg. 80). When controlling for attributional style, only gender, political ideology, and region of origin remained significant, suggesting what the authors called a "consensus" with regard to support for capital punishment in certain offender populations. Specifically, protected populations as outlined by the Supreme Court (i.e., individuals with ID and those who are mentally incompetent), in addition to offenders

described as "immature," were less likely to be sentenced to death. One inconsistency noted by Cochran and colleagues (2003), however, was that the latter of these elements, while dispositional in nature and generally used to support more punitive decisions, performed just the opposite, meaning that dispositional elements are not always used in that direction (i.e., further supporting the fact that some evidence is “converted” during capital mitigation). The researchers argued that while juror-level characteristics were important in understanding support for capital punishment, this effect was minimized when controlling for attributional style, indicating the relative strength and importance of this factor in understanding death penalty support and capital decision-making.

Juror-Level Characteristics

In addition to the content of mitigation and the themes by which information is presented, it is similarly important to understand juror-level factors that may influence how testimony is perceived by fact-finders. As mentioned previously, attitude constructs are more important in understanding support for the death penalty and/or receptivity to testimony relative to basic demographic variables, consistent with recommendations by Trahan (2011). Previous research has identified several characteristics of jurors that influence how evidence is interpreted and how decisions are made.

Death & life qualification. In *Witherspoon v. Illinois* (1968) the Supreme Court determined that potential jurors who would automatically and without discretion vote against the death penalty were ineligible to serve on a capital jury. This exclusion was later modified in *Wainwright v. Witt* (1985) to exclude potential jurors whose views, either pro- or anti-capital punishment, would “prevent or substantially impair” their ability to perform the duties of a capital juror. As such, only those individuals who fall in the middle of a continuum of death

penalty support (i.e., those who are neither so for or against capital punishment that they would always or never vote for it) are able to serve. From the outset, potential jurors are screened as to their fitness to serve on a capital jury; those individuals who meet the criteria are referred to as "death-qualified" while those who do not are referred to as "excludables." Butler (2008) suggested that the process of voir dire in capital jury selection rely upon attitude questionnaires designed to address related attitude constructs (e.g., authoritarianism, implicit bias) rather than relying upon stereotypic judgments derived from demographic characteristics of potential jurors.

In a sample of jury-eligible adults in California, Thompson, Cowan, Ellsworth, and Harrington (1984) found evidence that death-qualified individuals interpret evidence more favorably to the prosecution's side. Specifically, death-qualified participants demonstrated pro-prosecution sentiments across the four domains (i.e., credibility of witnesses, plausibility of facts, inferences from facts, attributions about the witnesses; Thompson et al., 1984). This finding provided evidence that juror attitudes, at least toward the death penalty, influence how evidence is interpreted which, in turn, influences ultimate decisions – an indirect relationship. The authors speculated that the differences may be the result of expectancy effects (i.e., death-qualified participants favor police control and prosecution and therefore expect and "fill-in" gaps with expected events), which is supported by differing responses on items about the inferences made from the ambiguous testimony.

Death-qualified subjects also expressed more regret for errors resulting in leniency (i.e., convicting an individual of a lesser offense or a finding of not guilty) and less regret for harsh errors (i.e., convicting an innocent individual or finding someone guilty of a more serious offense than what was committed) than excludable participants (Thompson et al., 1984). The authors stated that this relative difference suggests that death-qualified individuals are more

likely to convict and require less convincing to do so (Thompson et al., 1984). Thompson and colleagues noted that "death penalty attitudes may be superior predictors [of juror behavior and verdict] because they are associated with a set of beliefs about crime and criminal justice that provide available scripts for interpreting the complicated events of many criminal trials" (pg. 111). In other words, death penalty attitudes (i.e., those that render an individual death-qualified versus excludable) may be associated with a separate but related set of beliefs and/or behaviors that predict behavior more so than attitudes alone.

Luginbuhl and Middendorf (1988) similarly found that jurors who opposed the death penalty rated mitigating factors more highly in importance than did their counterparts who favored the death penalty; no difference was found for aggravating factors. The authors explained the finding in the following manner: "only those with strong opposition to the death penalty are willing to consider favorable evidence (or facts) that supports statutory or non-statutory mitigating circumstances and that points toward a more merciful sentence" (Luginbuhl & Middendorf, 1988, pg. 271). Similar findings were reported when comparing death-qualified jurors to excludable jurors – death-qualified participants endorsed aggravating factors more highly than excludable jurors; no differences existed between the two samples on the mitigating factor ratings (Luginbuhl & Middendorf, 1988).

In a vignette-based study, Butler and Moran (2007) found that death-qualified jurors endorsed more support for aggravating factors, less support for mitigating factors, and chose the death penalty more often than did excludable jurors. Death-qualified participants also scored higher on measures of legal authoritarianism, belief in a just world, and locus of control (indicating internal locus of control) than did excludable jurors. This study and the Vogel (2003) study also found that death-qualified jurors are demographically different - "men, Caucasians,

participants with lower levels of education, participants with conservative political beliefs, and participants with no prior jury service were more likely to favor the death penalty" relative to their counterparts (Butler & Moran, 2007, pg. 65). Vogel (2003) utilized a community sample of California residents and found that death penalty supporters were less liberally inclined, both religiously and politically, and of those who supported capital punishment, less than half (43%) supported LWOP as an alternative punishment. Income was the only significant predictor of LWOP such that as income increased, support for LWOP decreased.

In addition to death qualification, attitudes toward the death penalty have been shown to influence sentencing decisions and how one interprets evidence. In their review of the literature and interpretation of their 11 independent studies, O'Neil et al. (2004) determined that the relationship between death penalty attitudes and sentencing was most likely direct as attitudes directly influenced sentencing in that supporters were more likely to choose a sentence of death irrespective of other variables. In fact, "in every study, the effect of general support of the death penalty was greater than the effect of any manipulation of evidence or rating of aggravating or mitigating factors" (O'Neil et al., 2004, pg. 463).

Research on attitudes toward the death penalty has also demonstrated a relationship with considerations of mental illness. Specifically, "those high in support for the death penalty were also less likely to consider the defendant to be mentally ill (a variable sometime phrased as more likely to be in control of his or her actions at the time of the murder)" (O'Neil et al., 2004, pg. 462). This finding extended previous research which found that individuals in favor of the death penalty were less receptive to a mental illness defense, as evidenced by lower ratings of the likelihood that the crime was related to the defendant's schizophrenia (Poulson, Wuensch, Brown, & Braithwaite, 1997).

In another study of the influence of attitudes toward the death penalty, mock jurors in support of the death penalty were more likely "to believe that this murder was premeditated, that the defendant would be a future threat to society, and that the defendant deserved to die" (Goodman-Delahunty, Green, & Hsiao, 1998, pg. 265) than mock jurors who did not support the death penalty. The same participants who opposed the death penalty were also more likely to attribute the cause of the defendant's actions to external influences (i.e., emotional stress and alcohol/drug use) than mock jurors in support of the death penalty. Similar to previous research already mentioned, Goodman-Delahunty and colleagues (1998) found that mock jurors in favor of the death penalty favored conviction more so than those opposed to it. In summary, individuals who support the death penalty appear to interpret evidence differently than similar individuals who do not support the death penalty, and this level of support has been associated with decreased receptivity to mental illness defenses and/or specific mitigating factors.

Authoritarianism. As outlined by Narby, Cutler, and Moran (1993), "the authoritarian personality is further identified by a tendency to hold to conventional values; to submit to strong leadership…to act aggressively toward deviants and out-group members; and to believe in the righteousness of power and control" (pg. 34). In a meta-analysis of the effect of authoritarianism, Narby and colleagues (1993) found a significant, positive relationship between legal authoritarianism and verdict or conviction-proneness. The relationship between authoritarianism and verdict was moderated by other variables including sample and mode of transmission. Specifically, the relationship between authoritarian beliefs and verdict was stronger in studies with actual jurors (versus mock or student jurors) and with live or video trials (versus audio trials or transcripts; Narby et al., 1993). In a more recent meta-analysis, Devine and Caughlin (2014) found that authoritarianism demonstrated one of the strongest relationships (r =

.17) to dichotomous guilt judgments relative to other juror and defendant characteristics, second only to trust in the legal system. Further, “the relationship [was] manifestly stronger when dedicated measures of legal authoritarianism are used (e.g., the RLAQ)” (Devine & Caughlin, 2014, pg. 119).

In an earlier investigation of authoritarianism and deliberation, Bray and Noble (1978) found that mock jurors high in legal authoritarianism were more likely to vote for guilt, change their verdict during deliberations, and vote for longer sentences than those low in authoritarianism. Mock jurors low in authoritarianism demonstrated a leniency shift post-deliberation while the opposite was true for those mock jurors high in authoritarianism (Bray & Noble, 1978). Connors and Heaven (1987) found that authoritarianism was related to right-wing preference and conformity to authority.

Legal authoritarianism has also been associated with more severe punishment and sometimes more guilty verdicts than those low in authoritarianism (Bray & Noble, 1978; Davis, Bray, & Holt, 1977; Butler & Moran, 2007). In fact, Butler and Moran (2007) found that there was an association between legal authoritarianism and an increased emphasis on aggravators and a decreased emphasis on mitigators (Butler & Moran, 2007). In that same study, Butler and Moran (2007) found that authoritarian participants were also more likely to choose a sentence of death for the hypothetical defendant in a vignette. In another mock jury study, Butler (2007) found that individuals espousing authoritarian beliefs were more likely to choose a sentence of death than their counterparts. Further, these legal authoritarian beliefs were associated with demographic variables including ethnic group and SES.

Political conservatism. Previous research has indicated that political ideology influences the type of attributions one relies upon. For example, people who identify with more

conservative values tend to make more internal attributions than do individuals with more liberal ideologies (Zucker & Weiner, 1993; Pellegrino, Querolo, Monarrez, & Valenzuela, 1997; Weiner, Osborne, & Rudolph, 2010). Internal attributions, those where the cause of behaviors are attributed to characteristics of the individual rather than the environment, are associated with more punitive decisions.

In addition to the association between conservatism and attribution style, research has found an association between conservatism, death penalty sentences, and executions (McCann, 2008). Specifically, in an examination of state death penalty sentences and levels of conservatism, McCann (2008) found that conservatism was a significant predictor of death sentences and executions. Interestingly, there was an interaction between political conservatism and threat (operationally defined as violent crime rates and non-white population), such that more conservative states awarded more death sentences under conditions of high threat (a composite of homicide and violent crime rates, and the minority population percentage of the state), while more liberal states awarded fewer sentences of death under those same conditions (McCann, 2008). This finding was consistent with authoritarian dynamic theory (Stenner, 2005, as cited in McCann, 2008) which posits that individuals high in authoritarianism become more authoritarian under conditions of threat, whereas less authoritarian individuals do the opposite. Similarly, Butler (2007) found that individuals with conservative beliefs were more likely to choose a sentence of death for hypothetical capital defendants than individuals with liberal ideologies, suggesting a relationship between conservatism and sentence, in addition to the relationship with attribution style and threat (McCann, 2008).

The Current Study & Design

The goal of the current study was to evaluate how two of the diagnostic prongs of ID, deficits in IQ and AB, are interpreted by jurors during capital mitigation. This research is relevant in the post-*Atkins* landscape whereby defendants who have been found guilty and failed to prove ID are in a position to present mitigating evidence in order to argue for a sentence less than death. The current study also examined the potential influence of juror characteristics (i.e., political ideology, authoritarianism, death penalty attitudes) on perceptions of evidence and ultimate sentencing decisions. It is unknown at this time how these attitudes, particularly those about the death penalty, influence considerations of ID and whether the impact is similar to that of perceptions of mental illness.

The current study utilized a two-factor between-subjects design. Jury-eligible and death-qualified individuals were recruited to serve as members of mock juries ranging in size from four to six participants. Participants were exposed to testimony by a defense-hired psychologist in a capital mitigation hearing whereby the defendant's deficits in intelligence and adaptive behavior were manipulated with regard to whether the diagnostic prong was met (i.e., either the diagnostic prong was satisfied or unsatisfied). The primary dependent measures for the current research were pre- and post-deliberation sentences (life in prison without the possibility of parole versus death), along with continuous and categorical ratings of various mitigating (and potentially aggravating) factors.

Hypotheses

Hypothesis one. Consistent with attribution theory (Weiner, 1985) and research on the backfire effect (Barnett et al., 2004; Barnett et al., 2007) which indicates that factors within the perceived control of an individual, particularly without an internal attribution, are perceived as

aggravating, it was hypothesized that ratings of a defendant's adaptive behavior deficits would be rated as more mitigating in the presence of low IQ, an internal attribution. Specifically, it was expected that the ratings of the mitigating value of the adaptive behavior deficits factor would interact with IQ such that participants would rate the factor as less mitigating in the absence of IQ deficits than when IQ deficits were present.

Hypothesis two. Outside of an undergraduate honor's thesis which revealed inconsistent and unexpected findings with regard to perceived culpability (perceptions of culpability differed between IQ scores of 67 and 73, with perceptions of a hypothetical individual with an IQ score of 78 sandwiched between the two; Brewer, Salekin, & Wood, 2014), previous research has not yet investigated the influence of IQ with regard to its mitigating value. In the current research, it was hypothesized that an IQ of 73 would be perceived as more mitigating than an IQ of 81.

Hypothesis three. Although previous research has indicated that ID is mitigating (e.g., Barnett et al., 2004; Barnett et al., 2007; Najdowski et al., 2009), previous research has always included the diagnostic label which may account for this difference. This is supported by the fact that research has indicated that individuals misattribute symptoms of ID when the diagnostic label is not provided (Scior, Addai-Davis, Kenyon, & Sheridan, 2013), and that general literacy of the term "ID" is low (Scior, 2011). Because a diagnostic label was purposefully not included in the stimulus materials for the current project and only descriptions of the diagnostic prongs were provided, no difference was expected between the condition in which the defendant would technically qualify for a diagnosis of ID and the remaining three conditions in which the defendant would not. In other words, it was hypothesized that participants would perceive the individual who technically qualified for a diagnosis of ID as equally culpable as the remaining hypothetical individuals who would not qualify for such a diagnosis.

Hypothesis four. Average ratings of the mitigating factors were expected to mediate the relationship between participant variables including legal authoritarianism, death penalty attitudes, and political conservatism, and post-deliberation sentence, consistent with previous research indicating an indirect relationship between attitudes toward the death penalty and sentence (e.g., Thompson et al., 1984; Luginbuhl & Middendorf, 1988; Butler & Moran, 2007).

Hypothesis five. Based on previous research that has indicated more punitive judgments occur in the face of internal, controllable, and stable attributions (Weiner, 2006; Graham et al., 1997; Shaver, 1985), certain mitigating factors are expected to be more mitigating than others. Specifically, attribution theory would predict that external, uncontrollable, and unstable factors such as those related to the environment (i.e., home environment and neighborhood, exposure to crime, violence, and guns, negative male influences) will be perceived as more mitigating or less associated with punitive ideas than factors that are internal, controllable, or stable (e.g., illicit substance use, childhood onset psychological disorder).

2. METHODOLOGY

Participants

Inclusion Criteria

Participants were recruited from the undergraduate psychology subject pool at a large southeastern university. Inclusion criteria for participation were adapted from Alabama Code – Section 12-16-60: Qualifications of Jurors in order to have participants closely resemble triers of fact in this state. For this study, the 12 month residency requirement was removed as this was likely to severely limit the pool of eligible participants. Alabama's juror qualifications are similar to those of other states with the exception of the age requirement (i.e., in Alabama, you must be 19 years old to serve on a jury). Given that most states and the federal government allow 18-year-olds to serve on juries, 18-year-olds were deemed eligible to participate in the present study. The following criteria for participation were provided to students via the Sona Systems online portal and also reviewed at the start of each data collection session:

Eligible participants must:

1. *be a citizen of the United States, and age 18 or older;*
2. *be able to read, speak, and understand the English language;*
3. *be able to render satisfactory jury service; and*
4. *not be a convicted felon.*

Additional inclusion criteria was taken from case law that established the use of only death-qualified individuals in capital jury service. The question presented to eligible participants,

adapted after *Witherspoon v. Illinois* (1968), *Wainwright v. Witt* (1985), and previous research (Stevenson et al., 2010), was as follows:

Which of the following four response options best describes you?

1. *I would always vote for the death penalty in situations that allowed it.*
2. *Although I am in favor of the death penalty, I would consider the facts of a case in determining a sentence.*
3. *Although I have concerns about the death penalty, these concerns would not impair my ability to impose it in the appropriate circumstance.*
4. *I would never vote for the death penalty, even in situations in which the facts of the case called for it.*

Participants who chose either response option 2 or 3 were considered "death-qualified" and therefore eligible to participate. Individuals who indicated they would never impose the death penalty were excluded as those individuals were best captured in response option 4 above. Individuals who would indiscriminately choose the death penalty, those captured by response option 1, were also excluded. Given that this information was presented online and only those who met the criteria signed up for data collection sessions, it is unknown how many students were deemed ineligible. However, estimates from the general population suggest that approximately 15% of the population are not death qualified (Fitzgerald & Ellsworth, 1984). In addition, rates from the pilot studies (discussed subsequently) suggested between approximately 8% (pilot study 3) and 13% (pilot studies 1 and 2) of undergraduate students enrolled in psychology courses were not death qualified, while a large study with California undergraduate students found approximately 20% of the sample were "excludables" (Robinson, 1995). No participants were dismissed from the current study due to ineligibility; all participants identified

as death-qualified, as evidenced by choosing response option 2 or 3 above, and met the juror eligibility criteria listed above.

Demographic Characteristics of Participants

The final sample of participants included 271 participants after removing the 59 individuals who did not satisfactorily respond to the manipulation check questions (see Results section for details). According to a G*Power 3.1.7 (Faul, Erdfelder, Lang, & Buchner, 2007) analysis that was run prior to the start of the study, 195 participants were needed in order to obtain a moderate effect size (i.e., 0.3) and adequate power (i.e., .95). This was a conservative estimate, however, as the hierarchical structure of the data which necessitates multilevel modeling adds an error term that is unaccounted for by G*Power. As such, the final sample included an adequate number of participants according to this estimate.

Participants were divided into 4-, 5-, and 6-person mock juries (see subsequent section on mock jury details). The majority of participants were female (n = 178; 65.7%) and Caucasian (n = 206; 76%). Participants ranged in age from 18 to 34, with the majority of participants being between 18 and 19 years old (76.8%; M = 19.15; SD = 1.66). While minimal, there were participants from additional racial/ethnic groups (17.3% African American, 2.6% Hispanic, 1.5% Asian, and 2.2% other), with six participants identifying as multiracial, Indian, Native American, or Pacific Islander. The composition of the sample is consistent with the overall representation of individuals who comprise enrollment in undergraduate, lower-level, social science coursework.

Most participants were raised in the southeast region of the United States (n = 205; 75.6%), and had completed less than one year in college (M = .82, SD = 1.03, Range = 0-5). Self-reported high school and college cumulative grade point averages (GPAs) ranged from 2.7 to 4.9 (n = 267; M = 3.67, SD = .47) and 1.6 to 4.3 (n = 243; M = 3.23, SD = .54), respectively.

The majority of participants identified as Republican (*n* = 127; 46.9%) or Democrat (*n* = 52; 19.2%), with additional participants identifying as Independent (*n* = 8; 3%), Libertarian (*n* = 11; 4.1%), and politically unaffiliated (*n* = 71; 26.2%). On the 11-point scale ranging from liberal to conservative, participants generally identified as moderate (*M* = 6.64, *SD* = 2.16, Range = 1-11). Only one participant (.4%) reported having served on a jury in the past.

Seven participants (2.6%) reported having been the victim of a violent crime, and 34 participants (12.5%) reported having a close other as having been the victim of a violent crime in the past. Examples of crimes perpetuated against participants and/or their relatives included murder of a close friend or family member (*n* = 6), rape or sexual assault of a close relative or friend (*n* = 9), robbery of a close friend or relative (*n* = 2), assault of a friend or relative (*n* = 3), and four participants reported they were assaulted, mugged, robbed, or raped. It is important to note that several participants chose not to disclose details of the event(s) they experienced, and several participants endorsed multiple events.

Mock Jury Size and Demographic Characteristics

Data collection sessions were run with mock juries that ranged in size from 1 to 10 participants, depending on participant sign-up rates. Only data from mock juries of 4-, 5-, and 6-persons were retained for the current study where mock jury data were used (i.e., all data were retained for analyses involving individual pre-deliberation determinations). Although the size of an actual capital jury in most states is 12, research suggests that the differences between 12- and six-person mock juries are not substantial (Saks & Marti, 1997; Mize, Hannaford-Agor, & Waters, 2007; Devine, 2012). In their review of the literature, the Judicial Council of California (2004) cited both the advantages (primarily economic) and disadvantages (mentioned below) of

smaller juries. Of particular concern was the fact that the potential for a Type II error, or erroneous acquittal, is increased with smaller juries.

Previous research suggests that larger juries tend to have greater minority representation, deliberate for longer periods of time, and demonstrate increased recall (for a review, see Devine, 2012). Similarly, larger groups are more likely to not only have individuals favoring each side (i.e., verdict or sentence), but also less susceptible to conformity given the likelihood of other mock jurors favoring their position (Asch, 1956). In a meta-analysis of jury studies, Saks and Marti (1997) found that larger juries discussed the testimony with more accuracy and recalled more facts than the smaller juries.

In a study of the size and polling method of mock juries, Kerr and MacCoun (1985) found that juries composed of fewer members (i.e., 3 versus 6 versus 12) hung less often. In their study, mock juries were only given a period of ten minutes to deliberate in groups which is far from realistic in terms of the time it takes real juries to deliberate. The authors reported that "mock jury size did not affect the relative likelihood of conviction versus acquittal, a pattern obtained in nearly all previous work" (Kerr & MacCoun, 1985, pg. 354). Of relevance to the current research was the fact that differences between the six and 12 person mock juries were attributed to initial distributions of verdicts and "not to differences in their decision-making processes" (Kerr & MacCoun, 1985, pg. 356). Based on their findings, the difference between a mock jury of six versus 12 was not significant such that decision-making differed among the groups, but instead was attributed to the initial proportion of those favoring guilt.

Consistent with previous mock jury research with mock juries of less than 12 (e.g., Lynch & Haney, 2009; Collins, 2009), the High Court determined that the provision of a six-person jury versus a twelve-person jury does not violate a defendant's constitutional rights (*Williams v.*

Florida, 1970; *Colgrove v. Battin*, 1973). Later, in *Ballew v. Georgia* (1978), the High Court determined that a jury of six for non-petty crimes was the absolute minimum and anything less was considered unconstitutional. One year later, the Supreme Court determined that juries of six must reach a unanimous decision as anything less violates the 6th and 14th amendments of the Constitution (*Burch v. Louisiana*, 1979). As such, juries of six are appropriate in some cases and require unanimous decisions, similar to the current study.

In a review of the literature on groups in general, Cummings, Huber, and Arendt (1974) reported that "groups of six were found to be the most conducive to the emergence of both effective and efficient leadership" (pg. 461). Increasing group size has been associated with less consensus and less participation from individual group members but also more variety in ideas. Previous research has revealed a preference for groups of six such that groups of fewer (e.g., 4 or 5) do not organize or establish leadership effectively (Cummings et al., 1974; Ziller, 1957). Still others have found that group size is related to satisfaction (Slater, 1958) such that complaints about group size were less often presented in groups of two to four individuals than groups with six or seven.

Given the research on group decision-making and the Supreme Court's determination that juries composed of six did not violate a defendant's rights, the intention was to collect data only from mock juries composed of six. However, due to difficulties with recruitment (i.e., individuals signing up for sessions, but not arriving to participate) and time restraints, mock juries composed of between four and six participants were included. That said, a comparison between six-person mock juries and the group of juries composed of four and five revealed no significant difference on the dependent variable of interest (i.e., post-deliberation sentence; χ^2 (n = 1) = 2.68, p = .10, phi = .23).

Data were collected from seven four-person, seven five-person, and 35 six-person mock juries (i.e., 49 mock juries overall). Data from two mock juries were not analyzed as too few participants showed up to participate (i.e., on both occasions, only one participant showed up to the session). In cases where more than six participants showed up to a data collection session, the participants randomly assigned a number greater than six were moved to an alternate location (see Procedure section for details). As such, 46 participants (17%) were moved to an alternate location and therefore not included in the final analyses involving post-deliberation data.

On average, mock juries deliberated 29 minutes (*SD* = 5.86, Range = 10-44). The range of time spent deliberating is limited as mock juries were required to deliberate for a minimum of 30 minutes regardless of the unanimity of the initial decision. As such, 24 mock juries terminated as soon as 30 minutes had elapsed as a unanimous decision had been reached before that time frame; 12 mock juries went beyond the minimum allotted time limit, though often for just one to seven minutes. The longest time for deliberation was 44 minutes. Eight mock juries deliberated on the case for less than 30 minutes despite standard prompts from the researcher to continue discussing their decision making. Despite deliberating on the case for fewer than 30 minutes, all mock juries remained in the room and did not complete post-deliberation packets until 30 full minutes had elapsed. The majority of randomly-assigned forepersons (see Procedure section on details of how the foreperson was selected) were female (*n* = 31; 60.8%), Caucasian (*n* = 35; 68.6%), and had selected a sentence of life without parole (*n* = 31; 60.8%).

In the majority of mock juries, the first explicit shift of one of its members was toward life without parole (*n* = 24; 47.1%), and the majority of mock juries unanimously voted on a sentence of life without parole for the defendant (*n* = 33; 64.7%). Most of the mock juries took an evidence-driven approach to deliberation (*n* = 24; 47.1%), emphasizing the evidence that was

presented and discussing the mitigating factors as opposed to focusing solely on the outcome (i.e., sentence) for the defendant. Data on the proportion of females, proportion of Caucasian participants, and length of time for deliberation for the four-, five-, and six-person mock juries has been included in Table 1 below for reference.

Table 1
Characteristics of Mock Juries According to Size, M(SD)

	Length of Deliberation	Proportion of Female Participants	Proportion of Caucasian Participants
4-person	30.00(3.35)	.61(.41)	.68(.35)
5-person	26.33(7.23)	.60(.33)	.63(.36)
6-person	29.54(5.96)	.65(.22)	.73(.23)

Materials

Jury Instruction Literature

Previous research has consistently demonstrated that jurors have a poor understanding of judge's instructions (Severance & Loftus, 1982; Severance & Loftus, 1984; Severance, Green, & Loftus, 1984; Elwork, Alfini, & Sales, 1987; Smith & Haney, 2011). In fact, public policy has been developed in order to address the problems associated with poor comprehension (see, for example, Arizona's reforms which included the allowance of juror questions and simplified instructions; Yarnell, 2005). For the present study, Alabama's standard pattern jury instructions were simplified with the help of empirical research which has identified the importance of psycholinguistic and cognitive psychology principles, timing, and format in improving the comprehension of judge's instructions.

Elwork, Sales, and Alfini (1977) investigated the effectiveness of simplified pattern jury instructions in a sample of undergraduate students (study 1) and volunteers (study 2). The simplified instructions included modifications to the vocabulary (e.g., eliminating legal jargon; replacing abstract with concrete terms), grammar (e.g., using active versus passive voice), and

organization (e.g., incorporating alternative organization of the structure) of the original instructions. Results indicated that the simplified instructions produced better comprehension scores relative to the original instructions (Elwork et al., 1977). In the second experiment, participants watched a video of a civil trial and were instructed with either the original or simplified instructions either (1) before, (2) after, or (3) both before and after the trial video. As expected, the simplified instructions produced better comprehension scores. There were no differences in task performance between the groups who received the regular instructions versus no instructions at all (Elwork et al., 1977). In terms of the timing of instructions, participants demonstrated better factual knowledge of the case when the instructions were presented both before and after the presentation of the trial video.

Severance and Loftus (1984) investigated instructional comprehension by looking into actual questions asked by criminal and civil juries over a one-year period. Archival records provided by judges indicated that nearly one-quarter of juries sent questions to the judge during deliberations and most of these questions reflected a misunderstanding of important legal concepts (e.g., burden of proof and intent; Severance & Loftus, 1984). The researchers subsequently modified instructions consistent with Elwork, Sales, and Alfini's (1982) step-by-step guide to simplifying judge's instructions and determined that the simplified instructions resulted in improved comprehension relative to unmodified instructions and no instructions at all. Similarly, Severance et al. (1984) found that the recommended changes by Elwork et al. (1982) resulted in better comprehension, fewer errors, and improvements in the paraphrasing of instructions. Interestingly, the authors found that ex-jurors performed better than current jurors which the authors attributed to "the inverse relationship between experience and leniency"

(Severance et al., 1984, pg. 225) also seen in research with judges who convict more often than jurors, and the educational differences between their sample of current and ex-jurors.

Research into the format of presentation has also been investigated. Elwork et al. (1982) reported that supplementing oral instructions with written instructions resulted in more efficient deliberations (Forston, 1975), increased self-reported comprehension (despite absent demonstrated improvements; Heuer & Penrod, 1989), and improved comprehension relative to only oral or only written instruction (Prager, Deckelbaum, & Cutler, 1989). Similarly, Brewer, Harvey, and Semmler (2004) compared instructions that were presented in audio format, an audio-elaborated format, and an audio-visual format and compared the performance of law students and novices on comprehension. The audio-elaborated instructions included examples and a verbal flowchart of the process while the audio-visual format added visual depictions of instructional elements and the flowchart with animations in addition to the examples in the audio-elaborated condition (Brewer et al., 2004). Although absolute levels of performance were low regardless of condition, the audio-visual format produced the best results relative to the other two conditions in the novice group; there was no effect of condition on the law student sample. These results suggest that oral instructions supplemented with written instructions and visual depictions produce higher levels of comprehension relative to typical, standard presentations of juror instructions (i.e., audio only; Brewer et al., 2004; Elwork et al., 1982).

Principles of cognitive psychology have also been utilized to improve comprehension of instructions (Smith & Haney, 2011). In their investigation, Smith and Haney (2011) compared the original instructions with instructions simplified with the psycholinguistic principles outlined by Elwork et al. (1982) and a “pinpoint.” In the “pinpoint” model, instructions included “a situation model with concrete and case-specific examples” (Smith & Haney, 2011, pg. 344; e.g.,

“defendant” was replaced with the actual name of the defendant). Comprehension scores in the “pinpoint” and simplified instruction conditions were higher than scores in the original instruction condition. There was also an effect for definitional terms and comprehension of specific factors. The format of the instructions did not influence verdict choice (Smith & Haney, 2011). Given the extent of research supporting modifications to instructions to increase juror comprehension, Alabama pattern jury instructions were modified for the current study.

Jury instruction modifications. The original, unmodified instructions taken from Alabama Code were 1,883 words with a Flesch-Kincaid grade reading level of 14.7. The instructions took approximately nine minutes to read aloud in their entirety. These instructions were modified to include (1) more simplified language, (2) case-specific examples, and (3) the defendant’s name. An example of one of the modifications is provided below. In it, the original instructions read:

> *Ladies and gentlemen of the jury, in a separate proceeding the defendant has been found guilty of the capital offense of murder during the commission of robbery in the first degree. In this proceeding, you will not concern yourself with the questions of guilt but rather with punishment. We are about to begin the penalty phase proceeding. The duty before you is to advise the court as to what punishment should be imposed upon the defendant for the crime of capital murder – either life imprisonment without the possibility of parole or death.*

The instruction provided above was shortened to the following:

> *Mr. Michael Stevens has been found guilty of murder during a first degree robbery. You will decide the punishment for Mr. Stevens – either life in prison without the possibility of parole (LWOP) or death.*

The final, shortened version of the instructions was 317 words with a Flesch-Kincaid grade reading level of 9.8 (see Appendix A). Consistent with prior research, these instructions were presented in multiple formats and at multiple times. Specifically, instructions were read aloud to participants by the researcher twice, both before and after presentation of the evidence, and supplemented with a visual, elaborated PowerPoint presentation in which the weighing of factors was demonstrated via Smart Art graphics (see Appendix B).

Demographic Form

Participants were asked to report their age, gender, ethnicity, place of birth, place where raised, political affiliation and orientation, year in school, high school and current GPA, along with items related to their previous experience with jury service (see Appendix C). Political orientation was measured with an 11-point scale of left-right orientation based on previous research indicating this format is a valid measure of political views (Kroh, 2007). The decision to include a measure of political orientation is that previous research has indicated a relationship between conservatism and death sentences by state and individual (McCann, 2008; Jacobs & Carmichael, 2004). A qualitative measure of political affiliation (i.e., Democrat versus Republican) was also included.

Trial Transcript

The primary stimulus material for this study, the mitigation transcript, was adapted from a mitigation hearing in the southeastern region of the United States. The defendant was charged with murder during the commission of a robbery in the first degree, a capital offense in that particular state. The defendant was sentenced to death following a 10-2 vote of the jury. The testimony provided was that of a defense-hired expert who presented information regarding relevant risk and protective factors through a biopsychosocial framework.

Identifying details of the case were changed substantially. In addition to modifying the identifying content of the testimony, the mitigating factors were modified in order to make the case less clear (i.e., less mitigating and less in favor of a sentence of LWOP). The decision to do so was the result of the piloting of the transcript with undergraduate students (see subsequent section on the pilot studies) whereby little variation was found in sentence.

The modifications changed the 290 double-spaced page transcript to approximately 13 double-spaced pages of testimony without the manipulations described below (see Appendix D). The Flesch-Kincaid grade reading level of the testimony, without the manipulations, is 7.1. With the manipulations, the Flesch-Kincaid grade reading level is 8.5. The testimony presented information on the biopsychosocial risk and protective factors that the jury may consider in formulating their sentencing recommendation. The transcript was presented via audio recording and supplemented with a written copy for each mock juror. The decision to present in multiple formats was to reach the largest number of participants, particularly given that some individuals may have preferred to read along while others may have preferred to just listen to the testimony.

The testimony was recorded on standard computer software with an external microphone. Editing of the transcript was completed with Audacity, a free computer software program developed for editing audio files. The part of the expert was read by an assistant district attorney with experience speaking in the courtroom. The other portions of the testimony (i.e., the district and defense attorney) were read by two males with similar speech characteristics to the district attorney (i.e., similar in tone, volume, and lack of accent).

Experimental manipulation. The testimony included the presentation of multiple risk factors that may have impacted the defendant's developmental trajectory (e.g., father absence, childhood psychological disorder, etc.). Two of the diagnostic prongs for ID were manipulated

(i.e., IQ score and adaptive behavior). The manipulations were inserted at the beginning of the transcript as the first risk factors that were discussed. The decision to place risk factors at the start of the testimony was to capitalize on primacy effects. Research has shown that information presented early is better recalled (see Hurlstone, Hitch, & Baddeley, 2014, for coverage of this topic).

In addition to maximizing primacy effects, a brief reminder of the condition-specific manipulations appeared toward the end of the testimony (see Appendix E). Specifically, the attorney asked the expert to explicitly state if any of the factors uncovered during the mitigation investigation were inconsistent with a diagnosis of mild ID and if so, what they were. For example, in the deficits condition (i.e., both IQ and adaptive behavior deficits are present), the expert testified that information was consistent with a diagnosis of mild ID; however, it was not the job of the expert to evaluate the defendant for a diagnosis and therefore no opinion was provided. In conditions where deficits were not present, the expert testified that the IQ score, and/or lack of adaptive behavior deficits, were inconsistent with such a diagnosis. Providing this information toward the end of the transcript served as a reminder to participants and also capitalized on recency effects, or the idea that information presented recent temporally is recalled with greater accuracy (see Hurlstone et al., 2014 for coverage of the topic). In the transcript, particular emphasis was placed on the specific and detailed connections between the factors and the ramifications thereof, as suggested by Trahan (2011), to counter resistance to mental illness and mental retardation (ID) defenses.

Testimony related to the defendant's current and past intellectual functioning and the impact that this had on his developmental trajectory (see Appendix F) is a focal point of the present study. The defendant's prior full scale IQ and two index scores were very close to 70,

which is the upper boundary that separates people who may qualify for the diagnosis of ID from people who do not. The scores were within the standard error of measurement (SEM; i.e., 70 +/- 5), which meets the first prong of the diagnostic criteria in all clinical settings, and in all legal settings as set forth in *Hall v. Florida* (2014). In half of the conditions, information about Mr. Stevens's (the hypothetical defendant) IQ remained the same. In the other half, his IQ scores were adjusted so that they were outside of the acceptable range (i.e., above 75), thereby rendering him ineligible for the diagnosis[3]. Specifically, each score was adjusted by 8 points in the positive direction (i.e., Verbal IQ = 83, Performance IQ = 81, and Full Scale IQ = 81). These scores fall in the borderline range and are outside of the boundaries for ID, but are still over one standard deviation below the mean. An 8-point increase was chosen as this value is approximately one and a half times the SEM and raises the defendant's scores to the upper boundary of the borderline range, outside of the acceptable range for ID.

Testimony related to the defendant's adaptive functioning (i.e., the second prong of the diagnostic criteria for ID) was also included in the transcript (see Appendix G). In half of the conditions, this testimony was such that the defendant did not meet the adaptive functioning prong and in the other half he did. The deficits demonstrated by the defendant included behaviors in the conceptual domain such as difficulty managing money and following a schedule, along with educational and academic application difficulties.

In the deficits-present conditions, the expert focused on the importance of extra assistance and supervision in order for the defendant to function according to age- and culturally-appropriate standards. In the no-deficit condition, the expert testified that Mr. Stevens did take longer to learn these skills, but stated that he eventually learned to function independently. It was

[3] In all but the rarest of case, IQ scores higher than 75 render the diagnosis of ID inappropriate. However, in some cases, such IQ scores are permissible in the face of substantial deficits in adaptive behavior.

made clear that he did not demonstrate deficits to the level that would satisfy the second prong of the diagnostic criteria for ID. Presenting the lack of deficits in this manner (i.e., difficulty learning tasks and requiring additional instruction at the outset), is more consistent with such an individual than someone without deficits at all. In other words, presenting the hypothetical defendant as having no adaptive skill difficulties would be inconsistent with his IQ in either condition (i.e., it is unlikely that an individual with a score in either the borderline or extremely low range would demonstrate a complete lack of adaptive behavior deficits). Again, the purpose of the current study was to evaluate how defendants who may have brought an *Atkins* claim before the court are perceived during mitigation; a complete absence of deficits would be inconsistent in such a case.

In short, the first and second prongs of the diagnostic criteria for ID were altered in a manner consistent with the clinical opinion of the expert (see Table 2). These manipulations served to illuminate the influence of these factors during mitigation and the interaction between them. An investigation of this interaction is important given the likelihood that these factors almost invariably enter into the sentencing phase. The hypothetical defendant in Condition 1 meets the first two diagnostic prongs for ID (i.e., no information is inconsistent with ID).

Table 2
Manipulations by Transcript

Condition	Prong 1: Intellectual Functioning	Prong 2: Adaptive Functioning
1	Satisfied	Satisfied
2	Not Satisfied	Not Satisfied
3	Satisfied	Not Satisfied
4	Not Satisfied	Satisfied

Piloting the transcript. In order to establish the baseline ratio of life to death sentences and to ensure the strength of the manipulations, study materials were piloted with three groups of undergraduate students in a classroom setting presented as an educational activity (i.e., a lecture

followed both class activities tying the activity to class content). In accordance with institutional review board (IRB) requirements, students were provided with a comparable alternate activity to earn credit for the assignment. None of the students chose this alternative activity; all students participated in the piloting of the transcript.

Pilot study 1. Students enrolled in a summer Introduction to Psychology course were provided in-class activity points in exchange for their participation. Two versions of the proposed transcript were used – one in which the hypothetical capital defendant met both the first and second prong of the diagnostic criteria for ID (i.e., full scale IQ of 73 coupled with adaptive behavior deficits; Condition 1) and the other in which the hypothetical defendant met neither of the first two diagnostic prongs (i.e., a full scale IQ of 81 coupled with some adaptive behavior deficits, though not severe enough to meet the diagnostic criteria; Condition 2). Participants were read a description of the crime (see Appendix H) and modified jury instructions (see Appendix A; outlined in the materials section), and then asked to read one of the two transcripts and provide answers to the following questions:

1. If you were a juror in this case, which of the following sentences would you choose for Mr. Stevens?
    a. Life in prison without the possibility of parole
    b. Death
2. What was the defendant's IQ score?
3. Was Mr. Stevens' IQ score low enough to meet the first prong of intellectual disability?
4. Did the defendant demonstrate deficits in adaptive behavior, the second prong for intellectual disability?
5. Which of the following statements best describes you:

a. *I would always vote for the death penalty in situations that allowed it.*

b. *Although I am in favor of the death penalty, I would consider the facts of a case in determining a sentence.*

c. *Although I have concerns about the death penalty, these concerns would not impair my ability to impose it in the appropriate circumstance.*

d. *I would never vote for the death penalty, even in situations in which the facts of the case called for it.*

6. To your knowledge, are you eligible to serve on a jury (i.e., U.S. citizen, at least 18 years of age, no felony convictions)?

Of the 22 participants, 19 participants were both jury eligible (question 6) and death-qualified (i.e., chose response option b or c on question 5). Of these 19 participants, two chose a sentence of death (i.e., 11%). With regard to the IQ manipulation, all 22 participants answered question 2 accurately (i.e., responding 73 for condition 1 or 81 for condition 2) and 18 of the 22 participants accurately answered question 3. Three of the four errors on question 3 were participants in condition 2 who inaccurately reported the defendant met the first diagnostic prong; despite this, their responses on the open-ended question 2 were correct. With regard to the adaptive behavior manipulation, 15 of the 22 participants accurately responded to question 4. Five of the errors were made by participants in condition 2, and the remaining two were made by participants in condition 1. One possible explanation is that the errors were an artifact of the way the question was worded as one participant noted that “yes” the defendant had adaptive behavior deficits, though “not enough to be disabled.”

Table 3
Results from Pilot Study #1

	n	Death Sentence	IQ Errors	AB Errors
Condition 1	9	0	1	2
Condition 2	10	2	3	5
Combined Percentage	100%	11%	21.1%	36.8%

Pilot study 2. Given the low percentage of death sentences and number of errors related to the adaptive behavior manipulation, the transcript and question 4 were modified and piloted with a second group of undergraduate students. Modifications to the transcript included reducing the number and severity of the mitigating factors. For example, in the first version, the defendant's mother was diagnosed with major depression and was neglectful of her children. In the modified version, the defendant's mother worked several jobs in order to provide for her children. In addition, the family home was described as "small but tidy" rather than "dilapidated and dirty." These modifications were made to increase the percentage of death sentences to allow for more variability and sensitivity between conditions. In addition, it was important to establish that the adaptive behavior manipulation was strong enough to be detected by participants.

The modified transcript was piloted with 47 undergraduate students in a psychology course. Condition 1 (all deficits) and 2(no deficits) were again piloted in an effort to investigate the level of sensitivity between the two. Of the 47 participants, 87% were jury eligible and death-qualified (6 participants were excluded from subsequent analyses). Results of the piloted transcript are presented in the table below.

Table 4
Results from Pilot Study #2

	n	Life Sentences	Death Sentences	Percentage of Death Sentences
Condition 1	19	18	1	5.2
Condition 2	22	16	6	27.2
Combined Percentage	100%	82.9%	17.1%	--

Feedback was elicited from participants who chose life imprisonment without the possibility of parole with the following prompt: *Why did you choose life? What would make him death worthy?* The most frequent response was related to how low the defendant's IQ score was as this was mentioned by three participants. Other participants also noted that "his background was so bad" and suggested that the severity of the crime would need to increase in order to make him more worthy of a sentence of death. With regard to the manipulation checks, 100% of participants passed the IQ manipulation check and 94% passed the AB manipulation check.

Pilot study 3. Based on feedback from the second pilot of the transcript, the crime was modified in the following ways. First, the crime was modified such that the defendant, Mr. Stevens, acted alone and was therefore independently responsible. Second, the gender of the victim was modified in an effort to elicit possible sympathy and increase the likelihood that the victim would be viewed as vulnerable. Finally, the theft component was changed to stealing from the victim rather than a corporation, because some individuals are able to separate and/or justify stealing from a corporation due to the lack of an identified victim and the belief that no one is technically harmed (i.e., the "Robin Hood myth"). Elements which remained the same were that the victim continued to be random, the weapon was the same, and the component of robbery continued to be contained within the description.

The content of the transcript was also modified to reduce the salience and number of mitigating factors. Examples included modifying some of the mitigating factors to be more ambiguous (e.g., instead of growing up in a low socioeconomic home with an absent mother and abusive father, Mr. Stevens grew up in a middle class home and his mother was working a lot and therefore less present; further, the friends he "chose" were individuals in a lower economic area, which is where he was exposed to crime, guns, and drugs; his older brother was successful

in the sense that he was an athlete in high school and later while he was enrolled in college). The modifications were chosen in accordance with attribution theory (i.e., decreasing the salience of the factors via introducing some ambiguity as to whether the factor was within or outside of the defendant's control). Again, Conditions 1 and 2 were piloted with a group of undergraduate students (n = 52). Seven participants (7.7%) were excluded as they were not death qualified and jury eligible. Results from this pilot are included in the table below. Given the results from the second pilot study, the manipulation check questions were not included in this pilot study.

Table 5
Results from Pilot Study #3

	n	LWOP (*n*)	Death (*n*)	Percentage of Death
Condition 1	20	16	4	20.0
Condition 2	26	15	11	42.3
Combined Percentage	100%	67.4%	32.6%	--

Reflections from the pilot studies. Taken together, results of the pilot studies suggested that the manipulations in the transcript were both effective and salient enough to be detected by participants. In addition, results suggested that the mitigation presented in the initial transcript was too powerful such that participants were overwhelmingly sentencing the defendant to life rather than death. Variability and sensitivity were important to establish, particularly between conditions in order to detect effects. Modifications to the transcript and the crime allowed differences to be detected between Conditions 1 and 2, the deficits and no deficits conditions, respectively. Results of the final pilot suggested that participants were sentencing the defendant to death twice as much when no information was presented that was consistent with ID.

Pre-Deliberation Research Packet

Immediately after hearing the testimony, each participant was given a pre-deliberation research packet (see Appendix I). Participants rated each mitigating factor in terms of its influence and whether it factored into their sentencing determination (i.e., both a categorical and

quantitative determination of each factor). Moreover, prior to deliberation, each participant came to a conclusion about the sentence for the defendant (i.e., either a sentence of death or LWOP). This conclusion was made independently and served as a baseline to evaluate shifts post-deliberation.

The pre-deliberation packet also included a manipulation check related to the presentation of the transcript. The manipulation check allowed investigation as to whether or not the manipulation was strong enough and, therefore, able to produce reliable changes in the dependent variable(s) of interest. The manipulation check consisted of two yes-or-no questions related to the independent variables, IQ and AB deficits.

Post-Deliberation Research Packet

Participants completed a post-deliberation research packet after a unanimous decision was reached within their mock jury group (see Appendix J). This packet contained the following materials: (1) items similar to the pre-deliberation packet with regard to mitigating factor ratings, (2) the Revised Legal Authoritarianism Scale (RLAQ; Kravitz et al., 1993; see Appendix K), and (3) a measure of death penalty attitudes (O'Neil et al., 2004; see Appendix L). Support for the death penalty was measured via a 15-item, five-factor scale that was constructed and validated in 2004 (O'Neil, et al.). An instructional manipulation check (IMC) was included in the post-deliberation packet in order to identify those participants who did not pay adequate attention to the instructions. Recent research (Oppenheimer, Meyvis, & Davidenko, 2009) indicated that IMCs produced an increase in statistical power. The IMC directed participants to write "I read the instructions" somewhere on the page.

In addition to the IMC, an attention check question was also included. This question took the form of a single item embedded within a larger grouping of items related to death penalty

attitudes. The question read, *If you are reading this, circle number (5) below.* Participants who circled option (5) demonstrated their attentiveness to item content.

The Revised Legal Attitudes Questionnaire. The Revised Legal Attitudes Questionnaire (RLAQ; Kravitz et al., 1993; see Appendix K) was revised in 1993 and resulted in a 23-item measure of legal authoritarianism with three subscales: authoritarianism, antiauthoritarianism, and egalitarianism. The authoritarian and egalitarian subscales correlate with attitudes toward the death penalty (Kravitz et al., 1993). Research has indicated adequate reliability estimates (e.g., internal reliability: .71) in addition to adequate concurrent and construct validity (Kravitz et al., 1993). A revised 21-item version of the measure, translated to Spanish, indicated two- and three-factor models were better fits than a one-factor model (i.e., authoritarianism, antiauthoritarianism, and, in the three-factor model, egalitarianism; Garcia, de la Fuente Solana, Martinez Arias, & Fuente Sanchez, 2004). The decision to use the RLAQ is further supported by the results of a recent meta-analysis that revealed this measure demonstrated the strongest relationship to outcome measures relative to other measures of authoritarianism (Devine & Caughlin, 2014).

Death penalty attitudes questionnaire. Attitudes toward the death penalty are multifaceted, which is consistent with reasons cited by those who support (or do not support) the death penalty (e.g., vengeance versus fiscal; O'Neil et al., 2004; see Appendix L). In order to more closely investigate these attitudes, O'Neil et al. (2004) conducted a series of studies to identify and measure attitudes toward the death penalty across multiple facets and investigate the relationship of these attitudes toward decision-making.

The death penalty attitudes questionnaire (DPAQ) used in this study is a 15-item measure that was constructed and validated in 2004 (O'Neil et al.). The measure separates attitudes

toward the death penalty into five factors including general support and ideas regarding the deterrent and retributive nature of the penalty itself. This scale was used in 11 studies at the time of its publication and was subjected to stringent analysis of its psychometric properties (O'Neil et al., 2004). In short, convergent validity was demonstrated via relationships between the various factors and other related measures (e.g., dogmatism, punitiveness, and due process beliefs), and authoritarian beliefs correlated with scores on all five factors of death penalty support (O'Neil et al., 2004). There was evidence for mediation in that support for the death penalty was related to both perceptions of defendant dangerousness and verdict. Similar findings were found to support the fact that considerations of mental illness on part of the defendant mediated the relationship between death penalty support and ultimate verdict (O'Neil et al., 2004); though the authors noted that the direct effect of death penalty support remained large.

Procedure

Participant Sign-Up

Eligible students accessed the study's description and related information via SONA, the online subject pool portal for the Department of Psychology's research requirement. Interested and eligible students signed up online for one of the available three-hour long sessions; the description of their scheduled session provided information regarding the designated time and location for the data collection session. It was the responsibility of the student to either attend their scheduled session or cancel their participation in order to not receive a no-show penalty, consistent with subject pool guidelines.

Eligibility

Once all of the participants arrived to their designated session, the researcher reviewed the inclusion criteria (i.e., the criteria listed on the online experiment sign-up portal) to ensure

that all participants were both death-qualified and eligible for jury service, as previously outlined. Inclusion criteria was both read aloud to participants and supplemented with PowerPoint slides. Participants who did not meet inclusion criteria were to be dismissed and not receive credit nor given a penalty; all of the participants met eligibility criteria.

Information Statement and Random Assignment

At the start of the session, the researcher reviewed the information statement aloud with participants and answered any questions. Before materials were handed out to participants, each individual was instructed to put all of their belongings, including cell phones, electronic devices, and backpacks/purses, by the door of the room. This provision was to decrease the likelihood that participants would be distracted during the study or working on other tasks during data collection. Participants were then given a notecard with a number between one and twelve (the number of slots for each session) to place on the desk in front of them. This number served as their unique participant number for that data collection session. Numbers were distributed face down in random order to ensure random assignment to participant number. The place cards also ensured that data collected from each individual was correctly matched with other data sources provided during the session.

Following assignment to participant number, individuals were asked to complete a demographic form (Appendix C). Once demographic forms were completed and collected by the researcher, participants were informed that they were participating in a mock jury study. The researcher stressed the seriousness of the research and the potential to influence court cases in the future (Appendix M). The method of instilling a sense of the personal investment in the study was carried out in a manner similar to that of Najdowski et al. (2009).

Stimulus Material Presentation

First, a brief description of the elements of the crime were read aloud to participants (see Appendix H). Next, the modified jury instructions were read aloud to participants (see Appendix A) and supplemented with a PowerPoint presentation (see Appendix B). Participants were not permitted to ask questions about the instructions as this is not allowed in Alabama courts.

Following the presentation of the instructions, participants listened to the audio-recorded testimony and were provided with the opportunity to follow along on a hardcopy. Dual format presentation of material is thought to increase the chances that participants follow along and accurately understand the material (i.e., decreasing potential threats to internal validity). A PowerPoint slide, like those used in the actual re-sentencing hearing, was presented to participants during the audio recording of the testimony (see Appendix B). The slide listed the mitigating factors in the order they were presented during the testimony.

At the conclusion of the testimony, the researcher refreshed the jurors' memory for the modified jury instructions by reading them aloud a second time, again to increase the chance of adherence and comprehension of the task (Appendix A). Following this, pre-deliberation research packets were distributed to participants in the mock jury (Appendix J). Once all packets were completed, the deliberation phase of the study began for mock jurors.

Deliberation Phase

Once all of the pre-deliberation packets were collected, participants that had been assigned a number greater than six were excused to an alternative location with a research assistant. As previously mentioned, participants one through six served on the deliberating mock jury (see page 37 for rationale). The participants dismissed to the alternative location also

deliberated and participated in identical procedures, however, their data were not analyzed with the larger group because jury size was held constant (see page 37 for rationale).

Instructions for deliberation were read to participants by the researcher (see Appendix N). The participant who was randomly assigned the number one (or number seven for the alternate group) served as the foreperson during deliberations. For the purpose of future qualitative analyses, all mock jury deliberations were audio recorded and later transcribed. Each mock jury was provided one hour to reach a determination regarding sentence, but were permitted to end at any point earlier once they reached a unanimous verdict and had deliberated for a period of at least 30 minutes. This time period, though shorter than most real deliberations, is longer than those that have been used in other peer-reviewed, published jury deliberation studies (i.e., most studies allow mock juries only 30 minutes to deliberate; Devine et al., 2000; Devine, 2012). During deliberations, a researcher was in the room to record the order of mock juror contributions and to complete a qualitative deliberation form (see Appendix O). Although the researcher was in the room during deliberations, he or she did not interact with participants. Mock juries who did not deliberate or voted without discussion were instructed to discuss their decision-making (see Appendix P for standard prompts). The purpose of such an instruction was to encourage mock jurors to deliberate similarly to real world capital trials. A unanimous verdict was required based upon the decision in *Burch v. Louisiana* (1979) in which the Supreme Court determined that unanimity is required in cases with only six-person juries.

Following deliberations, each mock juror then completed a post-deliberation packet (Appendix J). In the case of the alternate group, if only one person was dismissed (i.e., seven participants arrived to the session), the individual was instructed to think about their decision for ten minutes prior to completing the post-deliberation packet. The decision to use ten minutes was

to allow participants the chance to consider their decision without over-burdening them with an excessive length of time to think without other participants in the room. Upon completion of the post-deliberation packet, participants signed out by providing their name and CWID, both of which were needed for participants to receive credit. Each participant was also provided with a debriefing form that included the contact information for the researcher and the Institutional Review Board should they have questions later.

3. RESULTS

Manipulation Checks

Strength of Manipulations

Two items in the pre-deliberation packet were designed to evaluate the strength of the manipulations, IQ and AB deficits. The questions were both yes/no questions and read as follows: *According to the expert who testified, did Michael Stevens have impaired intellectual functioning (i.e., IQ),* and *Did Michael Stevens demonstrate deficits in adaptive behavior severe enough to qualify for a diagnosis of intellectual disability (ID)?* The second question was revised following the pilot studies (see page 49 for details) in which incorrect responses were attributed to the way in which the question was asked. The first question, related to deficits in IQ, unfortunately remained the same as problems were not identified with the question during the pilot studies. Analysis of responses on these two items suggested a success rate of 55.6% (n = 184) for the IQ manipulation check and 82.2% (n = 272) for the AB manipulation check.

With respect to the IQ manipulation check question, there was evidence that accuracy depended upon condition as the rate of correct responses went from 25.6% to 89.7% when deficits were absent versus present, respectively. Examination of qualitative responses on a separate question further revealed that multiple participants (n = 24) accurately identified the hypothetical defendant's IQ despite failing the manipulation check. For example, five participants explicitly stated his true IQ score (i.e., 81) or the associated percentile, and 13 participants noted that his IQ was not low enough to qualify him as ID despite responding "yes" to the manipulation check question (i.e., that "yes" he had IQ deficits). Similar to the pilot study

results and subsequent modification of the AB question, inaccurate responses on this question were likely an artifact of the way the question was asked and not actual misinterpretation of the manipulation. Given the poor validity of the IQ question, participant data were not removed from subsequent analyses for their inaccurate response on the IQ manipulation check question.[4]

Conversely, rates of inaccurate responses on the AB question were comparable between the two sets of conditions with deficits versus no deficits (i.e., 18.9% and 16.0%, respectively). The 58 participants who inaccurately responded to this question were eliminated from final analyses, leaving the final sample with 272 participants, though data from the mock juries in which they participated were retained. The decision to retain the mock jury data was that although their misinterpretation could have contaminated the deliberation with misinformation, it is also possible that their misinterpretation was corrected via deliberations with their peers. Second, in the "real world" of juries, it is possible (and likely) that individuals serving on a jury will misinterpret or misunderstand information presented during the trial and subsequently "contaminate" the deliberations. As such, retaining the data from these "contaminated" mock juries was determined valid as similar occurrences and "contamination" is likely to occur in the real world of jury deliberations.[5]

Instructional Manipulation Check (IMC) and Attention Check Questions

As previously described, the IMC consisted of an instruction at the top of the post-deliberation packet instructing participants to write "I read the instructions" at the top of the page. It was designed to assess attention and used as an indication that they had read the

[4] A comparison between participants who failed versus passed the manipulation check revealed no significant difference with respect to the primary dependent variable, post-deliberations sentence (χ^2 = .298 and .017, p = .59 and .90 for conditions without and with IQ deficits, respectively).
[5] Further, a comparison between the mock juries with eliminated participants and those without revealed no significant differences on the dependent variable of interest (χ^2 = 1.50, p = .68).

instructions. Analysis of responses to this question indicated that 159 of the remaining 272 participants (i.e., 58.5%) failed this manipulation check, suggesting that over half of participants did not carefully read the instructions at the top of the post-deliberation packet. This percentage of failed responses is much larger than that reported in previous online research which has used IMCs (e.g., 92% and 89% success reported in Hauser & Schwarz, 2015; 90% in Goodman, Cryder, & Cheema, 2013).

Despite the failure of participants to accurately respond to this instruction, the instructions within each packet were arguably self-explanatory. As such, skipping instructions may not have altered the validity of the data. A comparison of participants who passed and failed the IMC indicated there was not a significant association with the dependent variable of interest, post-deliberation sentence ($\chi^2 = .12$, $p = .73$). In addition, the validity of the data, regardless of failure on the IMC, is further supported by the fact that only one participant missed the attention check question, suggesting that even though most participants were not carefully reading the instructions, they were paying attention to the content of the items to which they were responding. Elimination of the one participant who failed the attention check question brought the overall sample to 271 participants.

Descriptive Statistics

Pre- and Post-Deliberation Sentences

The primary dependent variables for this study were the pre- and post-deliberation sentences for the hypothetical defendant, Mr. Michael Stevens. The use of either pre- or post-deliberation sentence depended on the unit of analysis. Pre- and post-deliberation sentences for individual and mock jury units by condition have been included in Table 6 below. Visual inspection of the values in the table suggests a difference in sentence between Condition 2 (no

deficits) and the remaining three conditions as evidenced by a greater percentage of life sentences in conditions 1, 3, and 4 for both jurors and juries (i.e., approximately one-half versus two-thirds/three-quarters choosing a sentence of life). With the exception of Condition 4 where the number of people selecting a life sentence decreased by seven, pre- to post-deliberation sentences appeared relatively consistent and fluctuated by only two to three votes in either direction. In both conditions with IQ deficits, there was a leniency shift (i.e., an increase in life sentences following deliberation), while there was a shift toward death sentences in the remaining two conditions.

Table 6
Proportion of LWOP Sentences by Condition for Individuals and Mock Juries

	n_{juror}	Pre. Life[a] (*n*, %)	Post. Life[b] (*n*, %)	n_{jury}	Final Life[c] (*n*, %)
Condition 1	61	43, 70.5%	46, 75.4%	12	9, 75%
Condition 2	52	29, 55.8%	26, 50%	12	6, 50%
Condition 3	56	41, 73.2%	43, 76.8%	12	9, 75%
Condition 4	55	43, 78.2%	36, 65.5%	13	9, 69.2%

[a]Pre-deliberation LWOP sentences; [b]Post-deliberation LWOP sentences; [c]Final mock jury proportion of LWOP sentences

Ratings of Mitigating Factors

As a reminder, participants were asked to provide Likert-type scale ratings on a scale of one to ten for how mitigating and how aggravating each factor was, with higher ratings reflective of greater value of the factor. In addition, participants provided a qualitative determination of the mitigating value of each factor (i.e., each participant circled factors they perceived as mitigating). Ratings of the manipulated mitigating factors (i.e., IQ and AB) and the proportion of participants who endorsed each factor as mitigating have been included in Tables 7, 8, and 9. Only post-deliberation ratings have been reported here, as these values served as the primary dependent variable(s) in subsequent analyses.

Table 7
Perceptions of Mitigating/Aggravating Value of IQ/AB Deficits versus Not

Factor	*n*	Mitigating Rating *M*(*SD*)	Min.-Max.	Aggravating Rating *M*(*SD*)	Min.-Max.	Prop. Mitigating[1]
IQ (Deficits)	115	7.30(2.66)	1-10	2.88(1.97)	1-10	.77
IQ (No Deficits)	107	6.63(2.49)	1-10	3.07(2.00)	1-9	.66
AB (Deficits)	114	6.29(2.86)	1-10	3.11(2.10)	1-10	.65
AB (No Deficits)	108	5.73(2.63)	1-10	3.82(2.29)	1-10	.35

[1] Proportion of participants who identified this factor as mitigating.

Table 8
Perceptions of Mitigating/Aggravating Value of IQ by Condition

Condition	*n*	Mitigating Rating *M*(*SD*)	Min.-Max.	Aggravating Rating *M*(*SD*)	Min.-Max.	Prop. Mitigating[1]
1	59	7.07(3.01)	1-10	2.78(2.06)	1-10	.77
2	52	6.19(2.51)	1-10	3.38(2.15)	1-8	.65
3	56	7.55(2.24)	1-10	2.98(1.89)	1-9	.77
4	55	7.04(2.41)	1-10	2.76(1.81)	1-9	.67

[1] Proportion of participants who identified this factor as mitigating.

Table 9
Perceptions of Mitigating/Aggravating Value of AB by Condition

Condition	*n*	Mitigating Rating *M*(*SD*)	Min.-Max.	Aggravating Rating *M*(*SD*)	Min.-Max.	Prop. Mitigating[1]
1	59	5.95(3.10)	1-10	3.14(2.19)	1-10	.34
2	52	5.46(2.71)	1-10	3.52(2.21)	1-10	.39
3	56	5.98(2.56)	1-10	4.11(2.34)	1-10	.32
4	55	6.65(2.56)	1-10	3.09(2.01)	1-8	.64

[1] Proportion of participants who identified this factor as mitigating.

Review of Tables 7 and 8 indicate that participants found that information about IQ was mitigating, regardless of whether deficits were technically present or not. IQ was perceived as both more mitigating and less aggravating when deficits were present versus absent. Further, ratings of the mitigating value of IQ were highest in the condition whereby deficits in IQ were presented absent concomitant deficits in AB, while ratings of this factor were lowest in the condition where there were neither deficits in IQ nor AB. The mitigating value of IQ was relatively comparable in the remaining two conditions. The proportion of individuals who identified IQ deficits as mitigating was similar dependent upon whether deficits were present

(i.e., approximately 66% and 77% of participants rated IQ as mitigating when deficits were absent versus present, respectively).

A similar and seemingly stronger pattern emerged with respect to the mitigating rating of deficits in AB. As seen in Table 7, almost double the number of participants rated AB as mitigating when deficits were present relative to when they were technically absent (i.e., .35 to .65). The mitigating value of AB was highest in the condition whereby only AB deficits were present and lowest in the condition where neither AB nor IQ deficits were present. Ratings were almost identical in the remaining two conditions (i.e., both IQ and AB deficits were present; only IQ deficits were present). The proportion of individuals who identified AB as a mitigating factor was substantially higher in the condition where only AB deficits were presented relative to the remaining three conditions.

Ratings and qualitative determinations of the remaining mitigating factors and single aggravating factor have been included in Table 10. The range for each mitigating and aggravating factor rating was 1-10, therefore this column has not been included in the table. As seen from the ratings, illicit substance use, socioeconomic status, and exposure to guns were the lowest rated factors in terms of their mitigating value. The aggravating factor was the only factor that did not achieve an average mitigating rating greater than the midpoint of the scale, though the rating for illicit substance abuse was not far from the midpoint. In addition to falling close to the midpoint and therefore the least mitigating factor, illicit substance abuse was also the most aggravating factor. In addition, the lowest proportion of participants identified this factor as mitigating (i.e., .29). This is consistent with research on the backfire effect which has indicated that information about substance use is generally converted to aggravating information. Compared to the remaining factors in Table 10, IQ was the most mitigating factor presented in

the current study as evidenced by the highest ratings and the greatest proportion of participants identifying this factor as mitigating. The AB factor was relatively comparable to the other factors presented.

Table 10
Perceptions of Mitigating and Aggravating Factors

Factor	*n*	Mitigating Rating, *M*(*SD*)	Aggravating Rating *M*(*SD*)	Prop. Mitigating[1]
Childhood Onset ADHD	222	5.93(2.89)	3.05(1.98)	.50
Home Environment & Neighborhood	221	6.39(2.78)	4.00(2.47)	.61
Exposure to Crime & Violence	220	6.02(3.01)	5.44(3.03)	.52
Exposure to Guns	221	5.54(2.90)	5.53(3.00)	.42
Illicit Substance Use	222	5.19(2.82)	5.94(2.85)	.29
Lack of Adequate Parenting	222	6.37(2.63)	3.35(2.30)	.67[a]
Father Absence	223	6.39(2.74)	3.26(2.39)	.66
Negative Male Influences	222	6.43(2.81)	4.32(2.71)	.53
Failure to Provide Stability, Supervision, & Guidance	222	6.45(2.54)	3.63(2.32)	.51
Youthfulness	223	6.01(2.94)	3.85(2.49)	.46
Socioeconomic Status	223	5.46(2.64)	3.82(2.43)	.54[a]
Aggravating Factor – Crime was Committed During a Robbery	220	3.17(2.72)	8.00(2.61)	--

[1] Proportion of participants who identified this factor as mitigating.
[a]*Due to an error on the post-deliberation packets, the qualitative determination of these factors was not included. As such, the proportion here was calculated by looking to the proportion of participants who rated the factor a six or greater on the 10-point Likert-type scale as these values corresponded to a determination of more mitigating than not.*

Findings of Intellectual Disability

In addition to ratings of factors, participants were also asked to provide determinations as to whether they perceived the hypothetical defendant as demonstrating deficits commensurate with ID. As mentioned previously, Condition 1, where both IQ and AB deficits were present, was the only condition where the hypothetical defendant technically met the diagnostic criteria for ID. The third prong, though not manipulated in the current study, could be extrapolated from the testimony such that deficits in both domains (i.e., IQ and AB) were present during the developmental period. Participants provided a categorical rating as to whether they believed Mr.

Stevens had ID; these results are presented in Table 11 below. An exploratory analysis of the contribution of information about IQ and AB to these determinations has been included at the end of the results section. For now, however, it should be noted that the highest and lowest proportion of determinations of ID were in the expected conditions (i.e., the condition with the greatest and fewest deficits, respectively).

Table 11
Post-Deliberation Perceptions of ID by Condition

	Total *n*	Proportion of Findings of ID (*n*, Proportion)
Condition 1	61	49, .80
Condition 2	52	5, .10
Condition 3	53	23, .41
Condition 4	55	39, .71

Mediating Variables

Scores on the overall RLAQ-23 and its subscales reflect average ratings of the items composing that subscale; possible scores for each scale range from one to six. Scores on the Death Penalty Attitudes Questionnaire (DPAQ), on the other hand, reflect the sum of the items that comprise the subscale. As such, the scores presented in Table 12 are not comparable in terms of their absolute value. Descriptive statistics including means, standard deviations, minimum and maximum scores, along with Cronbach's alpha coefficients for each subscale have been included in Table 12. Using established interpretative guidelines from George and Mallery (2003), the reliability for two subscales was within an acceptable range (i.e., greater than .70); the remaining alpha coefficients ranged from .32 to .64. Scores on these measures did not vary by condition as evidenced by the results of a multivariate analysis of variance (MANOVA) and follow-up equivalence testing, consistent with methods from Weber and Popova (2012; $F(9,774) = 1.14$, $p = .32$, with significance values for between-subjects effects of .14, .69, and .49 for the RLAQ23, DP General Support, and Political Orientation scores, respectively; $p < .01$ for all comparisons).

Table 12
Descriptive Statistics for RLAQ-23 and DPAQ Subscale Scores

	No. of Items	*n*	*M*(*SD*)	Min.-Max.	Cronbach's alpha
RLAQ-23	23	220	3.65(.44)	2.35-4.70	.60
Authoritarian	8	222	3.55(.71)	1.38-5.00	.57
Anti-Authoritarian	6	223	5.83(4.25)	1.83-5.83	.43
Equalitarian	9	223	4.89(3.34)	1.33-4.89	.43
General Support	4	223	17.09(3.96)	6-29	.32
Retribution and Revenge	4	224	13.41(6.26)	4-33	.64
Death is Deterrent	3	224	14.92(3.95)	4-27	-.19
Death is Cheaper	2	224	11.65(4.78)	2-18	.85
LWOP Allows Parole	2	224	7.37(3.66)	2-18	.73

Hypothesis One

In order to evaluate the first hypothesis that ratings of AB deficits would depend on whether IQ deficits were present, a hierarchical linear model (HLM) was run with ratings of the mitigating value of AB deficits as the dependent variable, the IQ manipulation entered as the level 1 fixed effect, and jury's initial proportion of life sentences entered as the random-order level 2 effect. Only participants in conditions 1 and 4 were included in the present analysis, as the hypothesis related only to perceptions of the mitigating value of deficits in AB. Results of the unconstrained null effects model revealed that a significant proportion of the variance in the outcome variable (AB ratings) was accounted for by between-group variability, therefore providing justification for HLM ($\chi^2(23) = 63.73$, $p < .001$, ICC = .28). In other words, 28% of the variability in AB ratings is between groups (mock jurors), while the remaining 72% is between mock jurors within a given group.

The random intercepts model was run to investigate the influence of IQ in predicting ratings of AB deficits alone. IQ was entered as the level-1 variable; it was not a significant predictor ($\beta = -.61$, $SE = .80$, $t(80) = -.76$, $p = .45$). Next, a means-as-outcomes model was run to test the significance and direction of the relationship between the level-2 predictor and the dependent variable, ratings of AB deficits. The level-2 random effect was significant and

confirmed the relationship between the initial proportion of life sentences and ratings of AB deficits ($\beta = 4.96$, $SE = 1.57$, $t(22) = 3.17$, $p = .004$). In fact, the level-2 random effect accounted for 78.9% of the between groups variance in ratings of AB ($r^2 = .789$). In the final random intercepts and slopes mode, the level-1 fixed effect remained non-significant ($p = .39$), while the level-2 jury-level variable remained significant ($\beta = 4.90$, $SE = 1.62$, $t(22) = 3.02$, $p = .006$).

Hypothesis Two

In order to investigate the pattern of ratings on the IQ factor, a similar HLM model was run with IQ ratings as the dependent variable, the IQ and AB manipulations entered as level 1 fixed effects, and the jury's initial proportion of life sentences entered as the random level 2 effect to account for the nested structure of the data. Results of the unconstrained (null) model indicated that HLM was appropriate for the data as evidenced by a significant chi-square test ($\chi^2(48) = 149.92$, $p < .001$, ICC = .316). In other words, 31.6% of the variance in ratings of the IQ factor can be attributed to between-group variability, while approximately 68% can be attributed to variability within groups. Next, the random intercepts model was run to investigate the influence of the IQ variable alone in explaining the variability in the dependent variable. IQ was not a significant predictor; $p = .16$. Similar to the previous analysis, the means-as-outcomes model indicated that the level-2 predictor, initial proportion of life sentences within each mock jury, was a significant predictor of ratings of the IQ factor ($\beta = 4.49$, $SE = .69$, $t(47) = 6.51$, $p < .001$). In fact, this predictor explained 81.48% of the variability in ratings of the IQ factor. In the final random intercepts and slopes model, only the level-2 predictor was significant ($\beta = 4.42$, $SE = .77$, $t(47) = 5.77$, $p <.001$).

Hypothesis Three

In order to evaluate whether an unstated diagnosis of ID influenced perceptions of the defendant, a chi-square test of independence was run. As a reminder, it was expected that participants would view the individual as equally culpable (measured by final sentence) in the condition where he met the diagnostic criteria, but the diagnostic label was not provided (i.e., Condition 1). Given the nested structure of the data, pre-deliberation sentences were used in this analysis as there was no way to account for the hierarchical nature of post-deliberation sentences. The association between pre-deliberation sentence and a dichotomous categorization of whether the defendant technically qualified for a diagnosis of ID (i.e., Condition 1 was compared to the remaining three conditions) was analyzed. A significant association between the dichotomous classification and sentence was not found ($\chi^2(1, n = 225) = .05$, $p = .87$; OR = 1.07). The log of the odds ratio (log(OR) = .029) was within the range suggested to establish equivalence between groups (i.e., below .41 for a strict cut-off, or up to .85 for a more liberal cut-off; Wellek, 2010). In other words, participants were no more likely to spare the defendant from a sentence of death in cases where he technically met the diagnostic criteria for ID than in cases where he did not.

Although there was no association between the conditions where he qualified and did not qualify for the diagnosis with respect to sentence, this does not necessarily mean that participants perceived the individual with ID to be equally culpable, but instead they may not have recognized the defendant as having ID. In order to more fully evaluate the association between diagnosis and culpability, a similar analysis was run with perceptions as to whether the defendant had ID (a dichotomous determination) and sentence. In other words, this analysis investigated whether the participant viewed the defendant as having ID and therefore perceived him to be less

culpable, the latter measured by pre-deliberation sentence (again, because the hierarchical structure of the data could not be accounted for in a categorical analysis).

This analysis was significant with a moderate effect size, $\chi^2(1, n = 223) = 22.41$, $p < .001$, phi = -.32, suggesting that participants who perceived Mr. Stevens as having ID perceived him to be less deserving of a sentence of death, consistent with the High Court's determination in *Atkins v. Virginia* (2002). Calculation of an odds ratio revealed that the hypothetical defendant was four times less likely to be sentenced to death if the participant perceived him to have deficits commensurate with ID relative to if he was not perceived as such (OR = 4.04).

Hypothesis Four

Average ratings of the mitigating factors were expected to mediate the relationship between participant variables (i.e., legal authoritarianism, death penalty attitudes, and political conservatism) and post-deliberation sentence; a hypothesis that is consistent with previous research (e.g., Thompson et al., 1984; Luginbuhl & Middendorf, 1988; Butler & Moran, 2007). Three separate logistic regression analyses were run with pre-deliberation sentence as the dependent variable, average rating of the eleven mitigating factors as the mediating variable (not including ratings of IQ/AB factors), and RLAQ23, DP General Support, and political conservatism scores as the predictor variables. Given the poor reliability of the DP General Support factor (i.e., .32) along with the heightened reliability and similar number of items on the Retribution and Revenge subscale, the latter was included in a separate model in place of the former.

The method for testing mediation was taken from MacKinnon and Dwyer's (1993) method for testing mediation with dichotomous outcomes, and follow-up testing with the Sobel test (as per Kenny, 2014 & Herr, 2006). The Sobel test was used to determine whether the

mediation models were significant. As an aside, the Sobel test measures mediation by comparing the effect of the independent variable with and without the mediator variable to determine if the change was significant. Mediation is generally tested in a four step process, parsing out the relationships between (1) the predictor and the outcome, (2) the predictor and the mediator, (3) the mediator and the outcome, and (4) the predictor and the outcome while controlling for the mediator. A general model of the relationships measured in mediation has been included in Figure 1 below.

Figure 1
Mediation Model Example

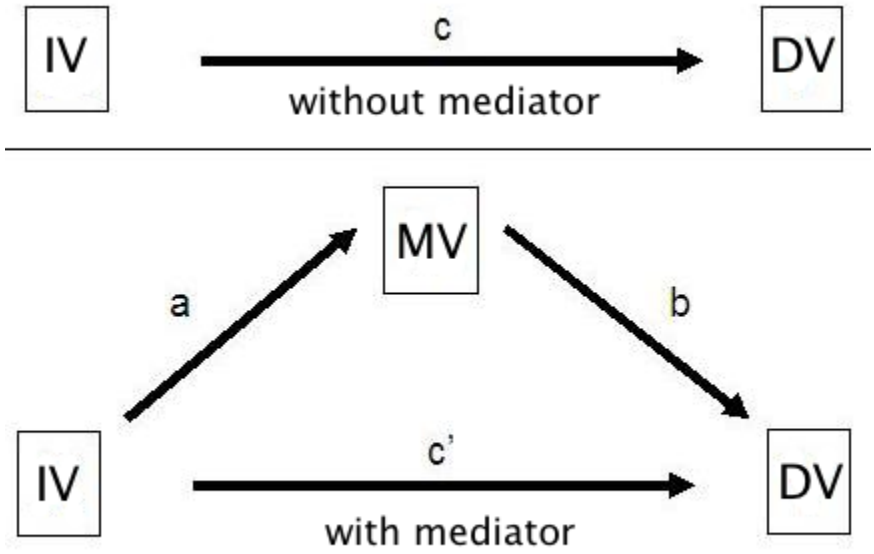


Table 13 includes the model coefficients for each of the mediation models. Per Kenny (2014) and Herr (2006), the absolute values are of greater importance than the statistical significance of those values. In the case of mediation, the coefficient for c' should be smaller than c, because the mediator explains part of that relationship therefore lessening the effect. When all of the relationship is explained by mediation (c'), this is referred to as complete mediation; partial mediation is when the value decreases, but does not reach zero.

As seen from the values provided in the table and the results of the Sobel tests, the relationship between pre-deliberation sentence and RLAQ23/political orientation scores was mediated by participants' average rating of the mitigating factors. In other words, political orientation and legal authoritarianism predicted individuals' acceptance of mitigating factors which, in turn, further predicted ultimate sentencing determinations for the hypothetical defendant. The relationship between pre-deliberation sentence and support for the death penalty was not mediated by mitigating factor ratings, and the relationship between Retribution and Revenge and pre-deliberation sentence was marginally significant.

Table 13
Mediation Model: Path Coefficients & Sobel Test Statistics

	Political Orientation	DP General Support	RLAQ23	Retribution & Revenge
a (IV-MV)	-.22	-.11	-.14	-.14
b (MV-DV)	-.57	-.60	-.61	-.62
c' (IV-DV, with M)	.21	.20	.08	.13
c (IV-DV)	.36	-.99	.18	-.99
Sobel [*z*, *p*]	$z = 3.26$, $p < .001$	$z = 1.62$. $p = .11$	$z = 2.14$, $p = .03$	$z = 1.92$, $p = .06$

Hypothesis Five

Based on previous research that has indicated more punitive judgments in the face of internal, controllable, and stable attributions (Weiner, 2006; Graham et al., 1997; Shaver, 1985), certain mitigating factors were expected to be more mitigating than others. Specifically, attribution theory would predict that external, uncontrollable, and unstable factors such as those related to the environment (i.e., home environment and neighborhood, exposure to crime, violence, and guns, negative male influences) would be perceived as more mitigating or less associated with punitive ideas than factors that are internal, controllable, or stable (e.g., illicit substance use, childhood onset psychological disorder).

In order to categorize the mitigating factors on the three dimensions, ratings from the primary investigator and four independent raters were solicited and Fleiss calculations, one method of establishing interrater reliability, were calculated for each dimension. In short, each rater was asked to rate the list of enumerated mitigating factors on three dimensions: locus of control, stability, and controllability. In order to establish interrater reliability, Fleiss calculations were made for each dimension. Fliess calculations look at the average correspondence between each pair of raters on the dimension, and then calculate an overall level of correspondence for that factor. Each rater is compared to each of the other raters, and the number of matches is calculated (i.e., the number of agreements among the total number of ratings). Then, an average correspondence rating is calculated for each pair of raters (i.e., Rater 1 & Rater 2, Rater 1 & Rater 3, etc.). Finally, an average level of correspondence is calculated by averaging the correspondence ratings for each pair of raters. Fleiss calculations for each dimension are provided in Table 14 below.

Table 14
Fleiss Calculations for Mitigating Factor Dimension Calculations

	Range of Correspondence	Average Correspondence
Locus of Control	.78-1.0	.88
Stability	.31-.89	.47
Controllability	.70-1.0	.84

As seen from the table, ratings for the Locus of Control and Controllability factors were generally consistent, ranging from .70 through 1.0, while ratings on the stability factor were unacceptably low (i.e., .47). Given the large discrepancy in ratings on the Stability factor, only ratings for the Locus of Control and Controllability factors were retained for this analysis. Next, mitigating factors were categorized according to each dimension. On the locus of control factor, there was 100% agreement on 10 of the 13 mitigating factors. For the remaining three factors,

three or four out of five raters agreed on the final rating. In those cases, the majority opinion was selected. On the controllability dimension, there was 100% agreement on nine of the 13 factors. The remaining had agreement among three or four of the five raters; majority opinion was again selected. The final ratings for each factor on these two dimensions is included in Table 15 below. Ratings where agreement was not 100% are denoted by asterisks.

Table 15
Attributional Dimensions for Each Mitigating Factor

	Locus of Control	Controllability
IQ	Internal	Uncontrollable
AB	Internal	Controllable*
ADHD	Internal	Controllable*
Exposure to Crime/Violence	External*	Uncontrollable
Substance Use	Internal*	Controllable*
Parenting	External*	Uncontrollable
Father Absence	External	Uncontrollable
Failure to Provide Supervision/Guidance	External	Uncontrollable
Youthfulness	Internal	Uncontrollable
Home Environment	External	Uncontrollable
Exposure to Guns	External	Uncontrollable
Negative Male Influence	External	Uncontrollable
SES	External	Uncontrollable

* *denotes categorizations for which there was not unanimous agreement.*

In order to test the influence of these dimensions on the ratings for each mitigating factor, average ratings for mitigating factors considered (1) internal, (2) external, (3) controllable, and (4) uncontrollable were calculated. These average ratings were entered into a two-way, within-subjects ANOVA, with Locus of Control (LOC) and Controllability as the within-subjects factors. Pre-deliberation ratings were used as the hierarchical structure was unaccounted for in this analysis. There were main effects for both LOC, $F(1,260) = 48.17$, $p < .001$, $\eta_p^2 = .16$, and Controllability, $F(1,260) = 6.75$, $p < .01$, $\eta_p^2 = .03$. Briefly, participants rated factors categorized as within the control of an individual as significantly less mitigating (i.e., controllable; $M = 5.73$, $SD = 2.03$) relative to mitigating factors categorized as outside of the control of an individual

(i.e., uncontrollable; M = 6.06, SD = 1.83). Similarly, mitigating factors rated as having an internal locus of control were rated as less mitigating (M = 6.01, SD = 1.83) than were mitigating factors rated as having an external locus of control (M = 6.05, SD = 1.93).

In order to more fully understand the interaction between the factors, follow-up analyses were run. Because there was not perfect correspondence between the factors (i.e., there were no factors categorized as external and controllable), only two specific comparisons were made. First, the effect of controllability was tested for factors considered to have an internal locus of control. In order to examine this, the average rating for internal, controllable factors (i.e., AB, ADHD; M = 6.03, SD = 2.13) was compared to the average rating for internal, uncontrollable factors (i.e., IQ, youthfulness; M = 6.47, SD = 2.13). This analysis was significant ($t(267) = -3.61$, $p < .001$, Cohen's $d = .22$), indicating that internal, uncontrollable factors are generally perceived as more mitigating than those that are internal and controllable. In order to examine the influence of locus of control on factors categorized as uncontrollable, factors with an internal LOC (i.e., IQ, youthfulness; M = 6.47, SD = 2.13) were compared to factors with an external LOC (i.e., father absence, SES, exposure to crime, parenting, poor supervision/guidance, home environment, exposure to guns, and negative male influences; M = 5.96, SD = 1.97). This analysis was also significant ($t(260) = 4.14$, $p < .001$, Cohen's $d = .25$), suggesting that when a factor is uncontrollable, it is more mitigating when it has an internal LOC relative to an external LOC.

Exploratory Analyses

Understanding Perceptions of Disability Status

As indicated previously (see results presented in Hypothesis Three), perceptions of disability status were important in sentencing as evidenced by a four times greater likelihood of receiving a sentence of LWOP when the hypothetical defendant was perceived as having ID. In

order to further understand how perceptions of disability status were made, a logistic regression analysis was run with the manipulated factors (IQ and AB deficits) as the categorical predictors, and perceived ID status as the dependent variable. This model was significant, $\chi^2(df = 2) = 98.70$, $p < .001$, with 76.7% of the cases correctly classified by the model. Further, the AB manipulation ($\beta = -2.55$, $p < .001$) was a stronger predictor relative to the IQ manipulation ($\beta = -1.28$, $p < .001$), suggesting a greater influence in determinations of disability status compared to information about IQ.

Understanding the General Influence of Presented Deficits

Although there was no control condition where information about IQ and AB were absent from the testimony, it was important to understand the influence of these factors when deficits were present versus not at all. The logic of this analysis followed from an examination of the descriptive statistics whereby differences in sentence were seemingly apparent between condition 2 (no deficits) and the remaining three conditions. A chi-square test of independence was run with pre-deliberation sentence and presence of deficits (i.e., condition 2 versus conditions 1, 3, and 4). The analysis was significant, $\chi^2 = 4.589$, $p = .03$, phi = -.13. Calculation of an odds ratio revealed that a sentence of LWOP was 1.88 times more likely when information about IQ/AB deficits was present versus absent.

4. DISCUSSION

The present study investigated the relative influence of deficits in intelligence and adaptive behavior during capital mitigation. In each case, the factors were presented as either (1) just within the range or (2) just outside of the range required to satisfy the first and second diagnostic prongs for ID. Investigating the influence of these factors is relevant for cases where a defendant's *Atkins* claim fails, and the defense team is subsequently afforded the opportunity to present mitigating factors to a judge or jury during capital mitigation. Understanding the relative influence of these factors when criteria are barely met and/or nearly missed can help in determining if and/or how to present this information to the trier of fact during sentencing.

The Influence of the Diagnostic Prongs

In the present study, the majority of individual mock jurors and juries decided on a sentence of LWOP for the hypothetical defendant, regardless of condition. The proportion of LWOP sentences was lowest when neither IQ nor AB deficits were present (i.e., approximately 50%), and relatively comparable among the remaining three conditions where deficits in one or both areas were present (between 66% and 78%). This suggests that evidence about one or both of the diagnostic prongs is influential with respect to sentence, as information about these deficits resulted in a substantial increase in the proportion of LWOP sentences relative to when deficits were not present. In fact, calculation of an odds ratio revealed that a sentence of LWOP was almost two times more likely in the conditions where deficits in either IQ or AB were presented relative to when both were technically absent.

There are two ways to interpret this finding: (1) deficits in IQ and AB are influential with respect to capital sentencing – doubling the chance of receiving a life sentence when present, and/or (2) information about "near misses" is harmful to a defendant, resulting in an increased likelihood of a sentence of death. With respect to the latter, and similar to interpretations from previous research (Gillespie et al., 2014), it may be that participants perceived the defense-hired psychologist as stretching aspects of the defendant's background in an attempt to allow the defendant to "shirk responsibility" for the offense on the basis of inappropriate factors when both factors were absent. In a sense, the combination of "near misses" in IQ and AB deficits backfired, resulting in a substantially greater likelihood of receiving a penalty of death. There were five explicit mentions of an IQ score of 81 in the 338 pages of transcribed deliberations; these excerpts (below) best demonstrate mock jurors' animosity related to the presentation of "near misses," particularly without accompanying deficits in AB.

> *Jury 11, Condition 2: Yeah, it was 81 and the average is 83 and the average ends at 85. It wasn't like it was like 60.*

> *Jury 4, Condition 2: His IQ was really low, and it wasn't that low but it was lower than average like still an 81 to 95. Even with an IQ of 85, that's still like the low end of average, like the low end. Like that's people who could barely function in regular classes, and he's even …and even if he's below that, then he's not bad enough to be in special education classes, so I think, that's a really harmful….and I've read a lot about that middle range and how, you don't really need special education and you aren't mentally disabled, but you aren't in the average range and so you are just…you are at a lower functioning ability, but you aren't disabled…. So are we going to say that's an aggravating factor because he wasn't actually disabled?*

> *Jury 52, Condition 2: Well, Forest Gump had an IQ of 48. He had an IQ of 81…so I don't see how a low IQ would make you want to kill someone. I feel like it doesn't at all.*

> *Jury 61, Condition 4: I think that IQ is a really... It's like they said "oh well technically because his IQ was an 81, he technically doesn't have a deficiency." But then they listed the categories and you only had to have one and he had at least two. [There was a] strong case for two [areas of deficits in adaptive functioning].*

> *Jury 27, Condition 4: It's also the mom's fault why his IQ is low…and it's not as if…they said and made it seem like…they said IQ from like 85 to 100 is normal, but his IQ was 81. They made it seem like his IQ was like 33 and it was like so low and he was so dumb…if he had studied a little bit more…it was like right there…if he knew a little more things…*

As a reminder, Condition 2 had neither IQ nor AB deficits, while Condition 4 had only AB deficits. As seen from the excerpts from Condition 2, participants responded relatively poorly to the presentation of an IQ of 81 as evidenced by a comparison to Forrest Gump, presumably in an effort to note that an IQ score similar to that presented is "not that bad" and "not like it was [a] 60;" one participant even suggested his IQ of 81 should be considered an aggravating factor because it did not technically fall within the range for disability. On the other hand, when an IQ of 81 was paired with concomitant deficits in AB, mock jurors noted the "strong case" for his behavioral deficits despite his IQ being outside the range. In the other example, the individual blamed the defendant's mother for his IQ, and subsequently noted that "if he had studied a little bit more," perhaps things would have been different.

These excerpts provide evidence for the latter interpretation – that the combination of "near misses" backfired and ultimately harmed the defendant's case. As seen here, presenting marginal deficits (i.e., an IQ of 81) absent accompanying behavioral deficiencies (i.e., deficits in AB) may backfire and subsequently be converted to aggravating information, similar to that seen in previous research with information about alcohol and drug use. As such, the explicit link between deficits, especially when marginal or "near misses," and their functional impairment – similar to the important nexus between mental illness and behavior as discussed in Wolbransky's (2011) research – is imperative in capital mitigation to avoid the potential for converting this information to aggravation.

Despite the relationship between the manipulated factors with respect to ultimate sentence, the ratings of the mitigating value of the IQ and AB factors were not associated with the manipulation of either factor. Instead, the initial proportion of life sentences within each mock jury explained a significant proportion of the variability in these ratings. The fact that this jury-level factor substantially impacted individual-level variables (i.e., independently-provided ratings of mitigating factors), suggests that jury deliberation had an important and profound impact on perceptions of facts presented during capital mitigation. In other words, it appears that group-level processes significantly influenced individual-level perceptions of mitigating factors such that the manipulation of those factors no longer mattered. This particular finding is consistent with previous research which has indicated that initial distributions of verdicts (in this case, sentences) are responsible for predicting the largest amount of variability in ultimate verdict (see Devine, 2012 for a review). It appears a similar process happened here such that the mock jury's initial leanings explained a substantial amount of the variability in ratings of the mitigating factors, rendering the manipulation of those factors unimportant.

Although the differences on the ratings of the items were not dependent on condition, there did appear to be a general difference in the mitigating ratings between the factors, with IQ (deficits present; deficits absent) being perceived as more mitigating than AB in all cases except when AB deficits were presented alone. In other words, IQ appeared to be a stronger mitigating factor relative to AB in almost all cases as evidenced by higher ratings of the mitigating value of the factor and also the greater proportion of participants who identified the factor as mitigating. This finding has practical utility for individuals involved in the process of capital mitigation as presenting information related to a defendant's IQ deficits (even when outside the range for ID) appears to be more mitigating relative to deficits in AB. However, and as discussed below,

deficits in AB were more influential with respect to perceptions of ID, which was strongly associated with sentencing determinations. In sum, this evidence suggests that information about the diagnostic prongs of ID was mitigating, regardless of whether the diagnostic prong was technically met, though presenting "near misses" on both prongs appeared to backfire and hurt the defendant's case for a sentence of life without parole.

The Influence of ID – Perceived or Otherwise

Consistent with expectations, there was no difference in culpability, as measured by sentence, between the condition in which the hypothetical defendant met the criteria for ID (Condition 1) and the remaining three conditions. This a priori prediction arose from previous research which has indicated that ID is mitigating, though this research has generally presented the diagnostic label absent behavioral descriptions. In other words, the diagnostic label itself has been found to be mitigating as opposed to the actual deficits associated with the disability (e.g., Scior, 2011; Najdowski et al., 2009). The results here are consistent with the notion that the label itself is mitigating, not necessarily the diagnostic prongs which comprise the diagnosis.

Extending this research, there was a strong association between perceptions as to whether the defendant met diagnostic criteria for ID and culpability. This finding suggests that perceptions of ID are important with regard to determining punishment. The perception that individuals with ID should be treated differently with regard to punishment is consistent with the High Court's ruling in *Atkins* (2002) whereby the death penalty was deemed unconstitutional for offenders found to have ID. The present study suggests that participants do perceive individuals with ID as less morally culpable for an offense as evidenced by the decreased proportion of death sentences awarded in light of (perceived) ID status. Similar to previous research (see Scior, 2011 for a review), this finding is also consistent with the idea that individuals misunderstand what

constitutes ID as only perceptions of disability status, not actual disability status, influenced culpability.

When looking at the proportion of individuals who perceived the hypothetical defendant to demonstrate deficits commensurate with ID (see Table 11), the proportion was highest in Condition 1, which was expected given that both IQ and AB deficits were present and the defendant technically met the diagnostic criteria for ID. Similarly, the lowest proportion of perceptions of ID was in Condition 2, where neither IQ nor AB deficits were present. A surprising finding was that the second highest proportion of ID determinations was in Condition 4 where only AB deficits were present, suggesting that deficits in adaptive functioning were more strongly indicative of ID relative to deficits in IQ. While this finding is consistent with the current diagnostic conceptualization of ID where severity is determined based on adaptive deficits (APA, 2013), prior conceptualizations of the disability and its severity have centered on deficits in intelligence. In fact, Greenspan and Switzky (2003) described the initial conception of ID as adding in Prong 2 (deficits in AB) almost as an afterthought that was subsequently paid little attention until only recently. Consistent with their argument that social vulnerabilities should determine whether an individual qualifies for protections (as opposed to arbitrary diagnostic determinations), participants in the present study appeared to rely more heavily on this type of information in making their own determinations of ID. This presumption was confirmed in an exploratory analysis whereby information about AB was more influential in perceived disability status relative to information about IQ as evidenced by larger beta weights in a logistic regression model.

In sum, it appears that descriptions of deficits in functional behavior are more influential than deficits in IQ scores when making determinations as to whether the diagnostic criteria for

ID are met. This is consistent with previous research that has indicated that mitigating factors related to mental illness were more influential (and therefore more mitigating) when the explicit nexus was shown between the symptoms of mental illness and one's ability to conform their conduct or behavior (Wolbransky, 2011). Here, deficits in AB, described in behavioral terms with respect to tasks the hypothetical defendant could not perform on his own, served as a more influential factor in determining whether ID was an appropriate diagnosis. That said, information about IQ, regardless of whether the diagnostic prong was met, appeared to be a more powerful mitigating factor than deficits in AB as evidenced by the quantitative and qualitative ratings of each factor.

In line with these findings, individuals in a place to present mitigating factors to fact-finders in the courtroom may do well to assume that deficits in both IQ and AB are influential and generally effective with respect to their mitigating value. While IQ deficits were generally more mitigating than AB deficits, the results of the present study suggested that the latter are more influential in making determinations of ID – a factor ultimately associated with individual and group sentencing determinations. In cases where disability status is the ultimate issue (e.g., *Atkins*), a preference for information related to the defendant's adaptive functioning deficits may prove more worthwhile.

Interpretation of Mitigating Factors – A General Model

Consistent with predictions, average ratings of the mitigating factors mediated the relationship between participant variables (e.g., legal authoritarianism, political conservatism) and sentence for the hypothetical defendant. Previous research has indicated that particular characteristics of individuals are associated with more punitive attitudes and influence verdicts and/or sentences. In the current study, the associations between sentence and political

orientation, legal authoritarianism, and attitudes toward the death penalty (i.e., general support and beliefs about the retributive and vengeful nature of capital punishment) were investigated. In two cases (i.e., political orientation and legal authoritarianism), scores significantly predicted ratings of the mitigating factors which, in turn, significantly predicted ultimate sentence awarded. This finding suggests that participant characteristics influence decision-making such that they color the way in which evidence is perceived (i.e., either more or less mitigating). In other words, higher levels of political conservatism result in lower ratings of the mitigating value of factors, resulting in greater support for a sentence of death as opposed to LWOP. Similarly, individuals who espoused greater legally authoritative attitudes perceived mitigating factors to be less mitigating than did individuals who espoused fewer authoritative attitudes.

This finding could be useful to legal defense and/or prosecutorial teams during jury selection with respect to identifying those attitudes which may influence one's receptivity to mitigating evidence. Similar to Butler's (2008) initial recommendation, asking questions and/or administering questionnaires during jury selection should focus on measuring attitude constructs (e.g., legal authoritarianism) that have been found to be associated with one's receptivity during mitigation. The present sample was composed of all death-qualified participants who tend to score higher on measures of political conservatism and legal authoritarianism (e.g., Luginbuhl and Middendorf, 1988; Butler & Moran, 2007; Vogel, 2003), as such there was likely a restricted range with respect to these characteristics relative to the general population. The fact that a relationship was detected despite a restriction in range suggests this relationship is robust and will be replicable in other samples. As such, the use of demographic and attitude questionnaires would likely be an ideal target for defense and/or prosecutorial teams in selecting jurors who may favor their side and/or be sympathetic to their argument(s) during capital mitigation.

Further, the attitude measures utilized in the present study were brief, can be administered in large groups, and have been previously validated, suggesting their use in this context would be appropriate.

Although the final factor, attitudes toward the death penalty (i.e., general support and retribution), was not significant with respect to the proposed mediation model, this may be the result of the poor reliability of the General Support factor. Further, individuals in this sample were death-qualified, and therefore likely presented as a relatively homogeneous group with respect to attitudes toward capital punishment. Although this enhanced the ecological validity of the current study, it did likely result in a restricted range of scores with respect to attitudes toward capital punishment, therefore limiting the potential for detecting an effect. Alternatively, it may be that attitudes toward capital punishment are simply less influential relative to legal authoritarianism and/or political conservatism, especially in death-qualified individuals. While it may be useful to investigate this relationship further, the present study provided evidence that other attitudes are simply more influential in this particular context (i.e., capital mitigation).

The Utility of Attribution Theory in Capital Mitigation

Finally, and consistent with attribution theory, factors that were categorized as uncontrollable or external to the individual were rated as more mitigating than those that were categorized as controllable and internal, respectively. Generally speaking, mitigating factors outside of the control of the defendant may be seen as imposed on the hypothetical defendant, and therefore the result of circumstance (i.e., a victim of chance, so to speak). However, individuals may place blame upon an individual for choosing something (i.e., a factor perceived as controllable may be viewed as chosen), therefore resulting in lower ratings of its mitigating value. This pattern was particularly true for factors with an internal locus of control. For

example, internal and uncontrollable factors (i.e., IQ and youthfulness) were rated as more mitigating by individuals than internal and controllable factors (i.e., AB and ADHD). In this case, all of these factors were rated as having an internal locus of control, originating within the person or as a characteristic of him/herself, yet those perceived as outside of the control of the person were more mitigating than those within the purview of the individual. This general pattern is consistent with previous research on the backfire effect whereby information about substance abuse – a factor considered internal yet within the purview of the individual – has been converted to an aggravating factor and used as evidence against the defendant (e.g., Barnett et al., 2004, 2007; Gillespie et al., 2014). Though not technically aggravating in the present study, the finding that similar factors were less mitigating when perceived as controllable is consistent with this research. In fact, the perceived controllability of deficits in IQ versus AB deficits may explain the difference in how mitigating these factors were perceived in the present study (discussed previously).

Based on this information, it would be useful for those developing mitigation defenses for capital defendants to consider the controllability of factors when determining whether to present this information in favor of a sentence of LWOP. Factors internal to the individual may best be framed as uncontrollable and therefore outside the purview of the defendant. As such, describing factors in concrete, behavioral terms with respect to their controllability would likely work in favor of the defendant. However, over-explanation may result in an alternative form of back-firing, mentioned previously, whereby the fact-finder may perceive the evidence as being "stretched." Carefully outlining the development of factors, with an emphasis on their controllability (if the factor is uncontrollable), will likely result in favorable interpretations by death-qualified jurors during capital mitigation.

For factors categorized as uncontrollable, however, an opposite pattern emerged with respect to locus of control. Specifically, uncontrollable factors with an internal locus of control were more mitigating than similarly uncontrollable factors with an external locus of control (e.g., IQ versus environmental factors such as SES, father absence, or the home environment). There are several possible interpretations of this finding, including that participants viewed uncontrollable and external factors as less important and more peripheral to an individual, and therefore less inherently mitigating. This interpretation is consistent with the importance of explaining resilience during capital mitigation, as fact-finders may wonder why a particular individual turned out to be someone who committed a capital crime while similarly situated others did not (e.g., "well his brother turned out fine – how come he didn't?"). The relative influence of external and uncontrollable factors may unintentionally communicate information about a defendant's resilience, or lack thereof, therefore proving aggravating as opposed to mitigating. Further research is needed to understand this pattern, particularly given that in general, factors with an external locus of control appear to be more mitigating, except for those factors perceived as uncontrollable.

Unfortunately, the list of mitigating factors did not include factors which were categorized as controllable yet external, and therefore the resultant impact of a factor when categorized as external and controllable is unknown at present. An example of a factor that might be categorized in this fashion is peer group, as persons generally choose their friend group (i.e., controllable), a factor or characteristic that is external to the individual. Although some factors presented in the present study could potentially be viewed as controllable and external (e.g., the fact that the hypothetical defendant "chose" to spend his time with individuals from a lower SES, in more crime-ridden neighborhoods, and without adequate parental supervision or support), the

individuals who provided the dimensional ratings for the factors categorized them according to their own conceptualizations absent information presented in the transcript. This fact speaks to the main limitation of this analysis – the arbitrary nature by which categories were created and subsequently tested. It would have been more meaningful to have each participant rate the mitigating factors on the three dimensions in order to investigate the relationship between perceived controllability, stability, and locus of control and ratings of their respective mitigating value. However, this was not done and therefore categories were created from responses provided by independent raters. While this alternative method would have provided more useful information about the utility of attributional dimensions in perceptions of mitigating factors, useful information was gleaned from the present study in spite of the arbitrarily categorized factors. Future research should investigate this possibility, as mitigating factors may not be perceived similarly by fact-finders; this hypothesis is bolstered by previous research that has suggested that more conservative individuals tend to rely on more internal attributions.

Limitations

Though the results of this study have made some important contributions with regard to knowledge of how deficits commiserate with ID are interpreted by death-qualified mock jurors during capital sentencing, the current research was not without limitations that must be considered. The primary limitations of the current research can be organized into three broad domains: (1) design limitations, (2) characteristics of the sample, and (3) problems in interpreting the manipulation check results. Each domain will be discussed in depth below, with a particular emphasis on how these limitations may impact the generalizability/interpretability of the results and suggestions for addressing these concerns in future studies in this area.

Design Limitations

Several limitations with respect to the design of the study bear mentioning as they concern the ecological validity of the current research and therefore limit generalizability. Because the current research was the first of its kind to investigate the impact of two of the three diagnostic prongs of ID during capital mitigation, internal validity was emphasized in the planning stages at the expense of external validity. As such, interpretations can be made with confidence regarding the impact of these specific factors, though the study setting did not closely approximate the real-world setting in which capital mitigation occurs.

First, and perhaps most salient, was that participants in this study were exposed to testimony from a defense-hired psychologist without having served as the decision-makers during the guilt-phase of the trial. In the real world of capital mitigation, jurors first hear testimony and make a determination as to the criminal culpability of the defendant prior to hearing evidence presented during mitigation related to the moral culpability of the defendant. In participating during this phase of the trial, jurors enter sentencing with a schema of the defendant, facts of the crime salient on their minds, and more often than not, with determinations as to the preferred sentence for the defendant (Bowers, 1995). In the present study, participants were provided only with a brief description of a crime committed by the defendant (see Appendix H), and asked only to consider the sentence for Mr. Stevens without knowledge of the facts of the crime. In fact, several mock jurors made reference to the lack of information related to the crime and stated their desire for additional information about the crime (e.g., motive, remorse, etc.).

Although one could argue that information presented during the guilt-phase is prejudicial, particularly given that humans struggle to consider only relevant information as evidenced by the

failure of limiting instructions offered during the course of trials (e.g., instructions to ignore pretrial publicity or stricken comments made during testimony; see Lieberman & Arndt, 2000, for a review), real-world capital jurors are charged with making determinations during both the guilt- and sentencing-phases of a capital trial. With regard to the impact of guilt-phase information, some researchers have called for different jurors for each portion of bifurcated trials and/or the use of non-death-qualified jurors during the guilt-phase, as "death qualification systematically distorts the attitudes of the jury in a direction that discriminates against the defendant and undermines the protections of due process" (Fitzgerald & Ellsworth, 1984, pg. 48). Given the fact that participants in this study were not exposed to information that may have arisen during the guilt-phase and additional information about the index offense, mock jurors were instructed to only consider the information that was presented to them without explanations of details of the crime. Though intended to measure only the impact of mitigation testimony, qualitative statements made during deliberations suggested that some mock jurors were filling in the gaps with their own beliefs related to the motive and level of remorse expressed and felt by Mr. Stevens. The following excerpts best demonstrate this tendency of participants to impute beliefs about Mr. Stevens and the crime itself, and also their recognition that information was missing:

> *Jury 42, Condition 1: In comparing that to this, we are not arguing on whether or not he did it. We know he did it. We are arguing whether the motive was extremely heinous [and] did he do this because he wanted to murder a woman or was there all these exterior and interior factors that caused him to do this. And we all agreed it was the exterior and interior that motivated him, even though there was still the aggravating factor that he knew he was murdering the woman in order to get the money, not because he wanted to kill her.*
>
> *Jury 35, Condition 1: It almost seemed like it was premeditated to me.*
>
> *Jury 37, Condition 3: We don't know how much remorse he felt. That's part of what's missing…we don't know how he felt. That would definitely sway my opinion. If he was*

proud of it on the stands, I'd be like this guy is a monster. I guess it is part of what you deal with in these kinds of studies. You don't get the entire picture. Also, I wonder if this was the only psychologist they hired. Normally they hire another one to avoid a bias and we're only hearing from one.

In addition to the lack of guilt-phase information, elements of the actual testimony limit the generalizability of the results to actual capital cases. Examples of ill approximations to reality include the fact that the testimony was presented in audio-only format, there was only one expert who was hired by the defense, and the transcript, while standardized for the purpose of consistency among conditions, was read by the voice actors therefore highlighting the differences between written and spoken language. The decision to use only audio-recordings of the testimony was made to limit the noise resultant from characteristics of the expert such as appearance, confidence, likeability, smiling behavior, etc., as research has suggested these elements impact credibility (e.g., Cramer, Brodsky, & DeCoster, 2009; Nagle, Brodsky, & Weeter, 2014). While the impact of this "noise" was reduced, it did create an artificial situation whereby the genuineness of the situation may have been inadvertently reduced as well. Similarly, the differences between written and spoken language (e.g., the formality of it, unnatural breaks, lack of repetition), may have contributed to the artificial feel of the presented material.

In addition to the content of the testimony, actual capital sentencing hearings generally include more than just one defense-hired expert, though research has indicated that jurors struggle to reason through and accurately make decisions when presented with conflicting expert testimony (Bourdreau & McCubbins, 2009). Additional witnesses sometimes include that of victims, family members, close friends, and others who may argue for and/or against the imposition of a death sentence for the defendant. Again, given the goal to best isolate the influence of the diagnostic prongs of ID during mitigation, only testimony from one psychologist

was presented to reduce the burden on the participant to balance, weigh, and reconcile differences from multiple parties. While effective in reducing the cognitive load on participants, actual capital jurors are tasked with balancing the competing information and reconciling those differences with the other members of the jury. Had testimony from additional witnesses been presented, the results may have differed with respect to the strength of the argument for the case of sparing the life of Mr. Stevens.

Elements of the deliberation phase of the study also limit generalizability as the foreperson for each mock jury was randomly assigned as opposed to selected by the members comprising the group, the size of the mock juries was small and not fixed in number (i.e., groups ranged from four to six participants), and time constraints were imposed on the audio-recorded deliberations. In the real world, jury members most often select the foreperson for the group, and this individual is generally male, Caucasian, and tends to be more extroverted, more educated, and of a higher SES (Feller, 2010). It is possible that selection of a foreperson influences deliberations as the foreperson usually talks the most and may have been selected on the basis of their ability to direct and lead conversations among group members (Devine et al., 2001). Randomly assigning forepersons in each group may have unduly changed the dynamic of the mock juries as individuals who may have been selected by group members may have been foregone in favor or a less dominant or less characteristic foreperson. That said, deliberations did not appear to be significantly influenced by the forepersons selected, as most juries engaged in relatively egalitarian conversations with little emphasis on the relative importance of individual members (i.e., the foreperson did not appear to talk more or direct others more so in most mock juries).

Though the mock juries were smaller in number than those used in actual capital cases, previous research (discussed at length previously) has indicated that the differences between six- and 12-person juries are not substantial and appear to relate primarily to the representation of minority members, etc. (Saks & Marti, 1997; Mize et al., 2007; Devine, 2012). While mock juries would have ideally been fixed in size with six members only, this was not feasible given the reliability of those recruited. However, comparisons of the juries with four, five, and six members revealed no significant differences on the primary variables of interest, suggesting less of an impact (cited previously: χ^2 (n = 1) = 2.68, p = .10, phi = .23).

Mock jurors were also forced to deliberate for a minimum of 30 minutes and were recorded during their deliberations. All but one of the mock juries reached a unanimous decision prior to the passage of the time limit, and the remainder of the time was often spent, following standardized prompts from the researcher, discussing how they arrived at their decision and which elements were most important in that decision. While this instruction and the recording of the discussions allowed an in depth examination of how the factors were interpreted by individuals, it did artificially prime participants to discuss elements otherwise absent from discussions, and the audio-recorder may have increased demand characteristics such that individuals may have felt less comfortable expressing less than favorable views. In response to this expected limitation, participants were reminded prior to deliberations that recordings were deleted immediately following transcription and reminded that their identities would not be matched to any statements made during the course of the deliberation.

Sample Characteristics

While the sample was large enough in order to achieve adequate power and detect effects based upon estimates from G*Power (Faul et al., 2007), the sample was largely recruited as a

matter of convenience. The sample was primarily female and relatively homogenous with respect to age, education, and ethnicity. Again, being one of the first of its kind to investigate this topic, the decision to emphasize the size of the sample over its representativeness was made intentionally. In addition, previous research has suggested that samples of undergraduate students do not differ substantially from actual jury samples (Bornstein, 1999), suggesting similar findings may result in research with more representative samples of death-qualified, jury-eligible United States citizens.

Generally speaking, the composition of actual juries, particularly those for capital cases, is homogenous with respect to (1) attitudes toward the death penalty, (2) political ideology, and (3) basic demographics given the requirement of death qualification. While all participants in the current study were death eligible, young females were over-represented, and all participants had a similar level of educational attainment and resided in the same region of the United States. The homogeneity of the sample and restricted range on certain characteristics (i.e., age, education) made sample weights similarly unfeasible. Sample weights have been suggested in cases where particular groups are over represented in samples. However, in order to adequately weigh some cases more than others, there should be representation of those groups in the original sample in order to successfully weight them more heavily. In the present research, there were no middle-aged, Caucasian males from the professional workforce who were represented. As such, sample weights were inappropriate. Future research may wish to include more diverse groups in their sample – even if sampling weights may still be needed – to better generalize to different groups.

Interpretation of (and Inconsistency between) Manipulation Checks

The current study included three different forms of manipulation check to ensure participant fidelity to the task and measure the strength of the manipulations. While well-

intended, the inconsistency between the results of these checks, along with the relative failure of two of the items made interpretation a challenge and required a detailed examination of convergent data to support the decision to retain and/or remove participants on the basis of these items. As a direct result of the pilot studies, the AB deficit question was modified as the ambiguity of the factor appeared to interact with the unintended ambiguity of the question and raised concerns as to whether failure on the item represented what it was intended to represent. Given the associated success of the IQ check during the pilot studies, however, that item was retained in its original form without modification. In doing so, and thankfully resultant from an unrelated question, failure on the item was attributed to an artifact of the question's ambiguity and not poor comprehension on part of the participant(s). As such, participants were not removed on the basis of failure on this item, which was supported by qualitative mentions elsewhere, the success of the item during the pilot studies, the apparent interaction with condition, and the equivalence between participants who failed and passed the item. Because of this decision, 38.7% of the final sample used for analyses (n = 87) failed the IQ manipulation check question. Although this decision was supported by the evidence presented above, it is unclear at this time if participants remained in the sample that may have not adequately responded to the manipulation or may have otherwise failed a more appropriate manipulation check. In other words, it is unclear how many of those 87 participants truly missed the manipulation or simply responded inconsistently given the poor content of the question. Obviously, we can have more confidence in results whereby manipulation checks successfully weed out individuals who have not paid adequate attention to study materials. At present, the strength of the IQ manipulation is unknown from the particular manipulation check question, however, its success can be extrapolated from other sources (mentioned previously). Should future research extend and/or

attempt to replicate the current results, it will be important to successfully modify the manipulation check questions to more fully identify individuals that did not pick up on the study's manipulations.

More complicated and perhaps less clear cut, however, was the decision to eliminate the IMC as a means of separating participants on the basis of their level of attention to the task. As mentioned previously, the IMC was designed to identify participants who did not adequately read and respond to written instructions provided on the research packets. Specifically, the post-deliberation packet included an instruction at the top of the first page to write "I read the instructions" on the page in order to demonstrate their attention to the task. However, less than half of participants completed this task (42.4%; n = 95). This proportion of success is far below that presented in previous research (e.g., 92% and 89% in Hauser & Schwarz, 2009; 71-86% in preliminary study and 54-65% in Studies 1 and 2, Oppenheimer et al., 2009). Fortunately, and as discussed elsewhere, the instructions contained within the research packets were not imperative to completion of the packets as the tasks were generally intuitive. As such, failure to read the precise instructions contained in the packet(s) was unlikely to result in misunderstanding or poor comprehension.

Future Directions

Future research should be designed with these limitations in mind in order to more fully understand the impact of the manipulated factors in the present design. While it is likely unfeasible to address all of these limitations in the same study, it may be worthwhile to design subsequent investigations with an emphasis on more closely approximating the real world with respect to individual design variables. For example, an identical study conducted with live (or video recorded) testimony from multiple experts and/or witnesses with mock juries composed of

exactly twelve members might result in different findings with respect to the primary independent variables. Similarly, providing more detailed or in-depth information about the index offense and/or charging the mock jury with first determining whether the hypothetical defendant is criminally culpable may influence subsequent interpretations of and/or willingness to consider mitigating factors. As mentioned previously, the current study was designed with an emphasis on internal validity at the expense of generalizability. As such, the level of experimental control and standardization of procedures was high, suggesting that changes in and/or differences between conditions on the primary variables of interest can be attributed to the experimental manipulation. However, the level of confidence we can have in how similarly situated information may be perceived during an actual death penalty sentencing hearing is less clear. Future studies, if designed to emphasize external validity, will help to increase that confidence.

The results of the present investigation indicated that (1) perceptions of ID influence culpability in the expected direction, (2) deficits in IQ and AB are mitigating in the minds of mock jurors and result in an increased likelihood of a sentence of LWOP, (3) perceptions of mitigating factors mediate the relationship between attitudinal variables and ultimate sentence for a defendant, and (4) mitigating factors are interpreted consistent with attribution theory dimensions. Further, and consistent with previous research, aspects of the deliberation process accounted for a significant proportion of the variability in dependent scores, suggesting that including deliberation in future research is both important and appropriate. Failing to include a deliberation component not only lacks generalizability with respect to how actual jury decisions are made, but also misses a key facet of individual-level determinations, as demonstrated here. In fact, jury group accounted for nearly one-third of the variability in the primary dependent

variable(s) of interest, a large and impressive effect. Future research should seek to understand the actual processes by which the deliberative process influences individual-level determinations and outcomes, as this may serve to improve this type of decision-making in the future. In fact, a qualitative analysis of deliberations from the current study would provide a wealth of additional information in terms of how ultimate decisions were made regarding the hypothetical defendant, Mr. Michael Stevens.

In keeping with the suggestion for qualitative analyses of jury deliberations from the current study, looking to the content of jury deliberations would likely assist in clarifying the decision-making process, but also identifying any errors in perceptions and/or stereotyped ideas about individuals with ID, the process of capital mitigation, and/or the burden of proof, for example. The following excerpts are just two examples of how stereotyped ideas about individuals with ID found their way into deliberations, and may have ultimately influenced each of the mock jurors' interpretations of the evidence:

> *Jury 36, Condition 1: Oh, another thing that I was kind of thinking about, before I like read what the doctor said, but like, if you meet someone who has an intellectual disability it's, a lot of times they are very loving, which was surprising to me because…. They don't like have a sense of personal space, so they are always like hugging people, and that's kind of like what I was like, he doesn't act like that, at all, he kind of refuses people, and I don't know if that's because of the friends he hung out with, or if that's an actual thing.*

> *Jury 37, Condition 3: I actually wanted to hear the guy himself. The psychologist touched on it but people with intellectual disorders, you can hear it in their voice. I think that would've confirmed it for me because the psychologist did have doubts about his intellectual disability. He has the IQ but his adaptive disorders aren't there, it's not a true intellectual disability, he's just really dumb.*

While there is evidence that naiveté and gullibility are social vulnerabilities often seen in ID (e.g., APA, 2013; Greenspan & Switzky, 2003), there is no qualifying source that suggests that people with ID are "always like hugging people" or that you "can hear it in their voice." These are examples of misconceptions of ID, similar to those espoused in *ex parte Briseno* (2004)

whereby a Texas court developed their own set of evidentiary factors on which to evaluate a defendant seeking protections from *Atkins* (e.g., citing the fact that close others can "see" the disability from a young age). These particular excerpts suggest that misconceptions about the disability are being used as evidence in capital mitigation and subsequently influencing interpretations of the evidence and ultimate outcome for the defendant. It is important to understand these misconceptions and challenge them in order to potentially ameliorate their impact.

In addition to misconceptions about ID, the content of deliberations suggested that participants in the current study were conflating ID with part of the legal definition of insanity (i.e., not knowing the difference between right and wrong). Several participants made reference to the fact that even though the defendant was impaired with respect to IQ and AB, he still "knew the difference between right and wrong" and was therefore responsible. These misconceptions persisted despite comprehensive teaching by the defense-hired expert as to the characteristics of ID (which did not include moral reasoning deficits or insanity judgments) and clear, concise instructions regarding how to interpret mitigating information.

A similar finding was reported by Haney and colleagues (1994) following interviews with actual capital jurors. One juror commented of the process, "I think the bottom line was, at the time he was committing [the crimes], did he know what he was doing? Did he know right from wrong? That's the whole thing" (Haney et al., 1994, pg. 162). Similarly, one of the capital jurors who was interviewed converted an IQ in the range of deficits to an aggravating factor secondary to a poor relationship with the legal concept of insanity (i.e., "[The defendant] had an IQ of 70 or 75, I believe. Right on the borderline of retard[ed]. I mean, he's just above the line. But that, still, we didn't feel affected his ability to know right from wrong," Haney et al., 1994,

pg. 164). This would suggest that expert witnesses or defense/prosecution teams may wish to spend time educating juries as to the threshold and manner in which decisions should be made, as well as distinguishing between different legal decisions (e.g., insanity, competence, determinations of ID, etc.). Future research might investigate the utility of further educating jury members on the more specific and technical legal criteria for which they are required to make their decision (e.g., the difference between insanity and ID or the difference between moral responsibility and criminal culpability).

As mentioned previously, it is important to more fully and appropriately investigate the utility of attribution theory in understanding how mitigating factors are perceived during capital mitigation. Preliminary evidence from the current study was consistent with a priori hypotheses that attributional dimensions would account for differences in the mitigating value of the factors presented. However, the measurement of these dimensions was less than ideal and the factors were categorized somewhat arbitrarily (i.e., interrater reliability was established between independent raters for which most, but not all, of the ratings were unanimous; the raters had not been exposed to the transcript and therefore made judgments of the factors out of context; the ratings could have differed from those of participants). Future research should measure perceptions of mitigating factors along these dimensions (i.e., ratings of the controllability, locus, and stability of factors) and how these perceptions influence perceived value of the factor. Relying on perceptions of these dimensions, as opposed to ratings from people external to the project, is in line with results from the current study where perceptions of ID (not actual ID) influenced outcomes. Soliciting ratings on these dimensions will also help to understand how these ratings are made, particularly for factors which may not clearly fall on one side of the dimension (e.g., some of the factors in the current study could be perceived as controllable or

uncontrollable, depending on how the individual interpreted them in the context of information presented during the testimony). In short, future research should focus on elucidating the process by which factors are considered in the mind of mock and real jurors, and whether attribution theory truly accounts for these differences, consistent with the present results.

The results of the present study contributed to existing literature in that there is now evidence to support the presentation of deficits commensurate with ID during capital mitigation, especially when the diagnostic prongs would be sufficiently satisfied in one or both domains. Qualitative analysis of deliberations from simulated jury studies can help to further understand how these factors are explicitly interpreted and understood by mock jurors when faced with the task of ultimately deciding the penalty for a defendant. Until then, however, deficits in IQ and AB should be considered during capital mitigation, as both appear to be perceived as mitigating and therefore in favor of a sentence of LWOP. Similarly, individuals involved in the jury selection process may find it useful to rely on attitudinal questionnaires in order to more successfully select (or dismiss) jurors who may prove favorable to (or unfavorable to) their case.

REFERENCES

Ala. Code § 7-1-101 (1975).

Alabama Code - Section 13A-5-40: Capital Offenses (Supp. 1985).

American Association on Intellectual and Developmental Disabilities (2010). *Intellectual disability: Definition, Classification, and systems of support.* Washington, DC: Author.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied, 70,* 1-70.

Atkins v. Virginia, 536 U.S. 304 (2002). Retrieved from https://supreme.justia.com/cases/federal/us/536/304/case.html

Ballew v. Georgia, 435 U.S. 223 (1979). Retrieved from http://www.law.cornell.edu/supremecourt/text/441/130

Barnett, M. E., Brodsky, S. L., & Davis, C. M. (2004). When mitigation evidence makes a difference: Effects of psychological mitigating evidence on sentencing decisions in capital trials. *Behavioral Sciences and the Law, 22,* 751-770. doi:10.1002/bsl.591

Barnett, M. E., Brodsky, S. L., & Price, J. R. (2007). Differential impact of mitigating evidence in capital case sentencing. *Journal of Forensic Psychology Practice, 7,* 39-45. doi:10.1300/J158v07n01_04

Bjerregaard, B., Smith, M. D., Fogel, S. J., & Palacios, W. R. (2010). Alcohol and drug mitigation in capital murder trials: Implications for sentencing decisions. *Justice Quarterly, 27*(4), 517-537. doi:10.1080/07418820903051623

Blume, J. H., Johnson, S. L., Marcus, P., & Paavola, E. (2014). A tale of two (and possibly three) Atkins: Intellectual disability and capital punishment twelve years after the Supreme Court's creation of a categorical bar. *William & Mary Bill of Rights Journal, 23,* 393-414.

Blume, J. H., Johnson, S. L., & Seeds, C. (2009a.). Of *Atkins* and men: Deviations from clinical definitions of mental retardation in death penalty cases. *Cornell Journal of Law and Public Policy, 18,* 689-773. Retrieved 17 August 2014 from http://www.lawschool.cornell.edu/research/jlpp/upload/blume.pdf

Blume, J. H., Johnson, S. L., & Seeds, C. (2009b.). Symposium: An empirical look at *Atkins v. Virginia* and its application in capital cases. *Tennessee Law Review, 76,* 625-639. Retrieved from http://www.iapsych.com/iqmr/blume2009b.pdf on 17 August, 2014.

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

Bourdeau, C., & McCubbins, M. D. (2009, February). *Competition in the courtroom: When does expert testimony improve jurors' decisions?* 3rd Annual Conference on Empirical Legal Studies Papers. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1121025

Bottoms, B. L., Nysse-Carris, K. L., Harris, T., & Tyda, K. (2003). Jurors' perceptions of adolescent sexual assault victims who have intellectual disabilities. *Law and Human Behavior, 27,* 205-227.

Bowers, W. J. (1995). The capital jury project: Rationale, design, and preview of early findings. *Indiana Law Journal, 70,* 1043-1102.

Bray, R. M., & Noble, A. M. (1978). Authoritarianism and decisions of mock juries: Evidence of jury bias and group polarization. *Journal of Personality and Social Psychology, 36,* 1424-1430.

Brewer, H., Salekin, K. L., & Wood, M. E. (2014, April). *Juror perception of culpability and its impact on verdict and sentencing: Defendants with and without intellectual disability.* Poster presentation at the University of Alabama Undergraduate Research and Creativity Conference, Tuscaloosa, Alabama.

Brewer, N., Harvey, S., & Semmler, C. (2004). Improving Comprehension of Jury Instructions with Audio-Visual Presentation. *Applied Cognitive Psychology, 18,* 765-776. doi:10.1002/acp.1036

Bruininks, R. H., Thurlow, M., & Gilman, C. J. (1987). Adaptive behavior and mental retardation. *The Journal of Special Education, 21,* 69-88.

Burch v. Louisiana, 441 U.S. 130 (1979). Retrieved from http://www.law.cornell.edu/supremecourt/text/441/130

Butler, B. (2007). The role of death qualification in capital trials involving juvenile defendants. *Journal of Applied Social Psychology, 37,* 549-560.

Butler, B. (2008, May). Caveats of the death-qualified jury. *The Jury Expert,* 1-14.

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

Cochran, J. K., Boots, D. P., & Heide, K. M. (2003). Attribution styles and attitudes toward capital punishment for juveniles, the mentally incompetent, and the mentally retarded. *Justice Quarterly, 20,* 65-93.

Cohen J. (1988). *Statistical Analysis for the Behavioral Sciences* (2nd Ed.). Lawrence Erlbaum and Associates, Publishers: Hillsdale, NJ.

Colgrove v. Battin, 413 U.S. 149 (1973). Retrieved from https://supreme.justia.com/cases/federal/us/413/149/case.html

Collins, P. A. (2009). The effects of cognitive disability evidence on death penalty dispositions: An analysis post *Atkins v Virginia*. *Criminal Justice Studies, 22,* 17-38. doi:10.1080/14786010902796465

Connors, J., & Heaven, P. C. (1987). Authoritarianism and just world beliefs. *Journal of Social Psychology, 127,* 345-346.

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

Cummings, L. L., Huber, G. P., & Arendt, E. (1974). Effects of size and spatial arrangements on group decision making. *Academy Of Management Journal, 17,* 460-475. doi:10.2307/254650

Cunningham, M. D. (2010). *Evaluation for capital sentencing.* New York: Oxford University Press.

Davis, J. H., Bray, R. M., & Holt, R. W. (1977). The empirical study of decision processes in juries: A critical review. In J. L. Tapp and F. L. Levine (Eds.), *Law, justice and the individual in society: Psychological and legal issues* (pp. 326-361). New York: Holt.

DeMatteo, D., Marczyk, G., & Pich, M. (2007). A national survey of state legislation defining mental retardation: Implications for policy and practice after *Atkins. Behavioral Sciences and the Law, 25,* 781-802. doi:10.1002/bsl.777

Devine, D. J. (2012). *Jury decision making: The state of the science.* New York: New York University Press.

Devine, D. J., & Caughlin, D. E. (2014). Do they matter? A meta-analytic investigation of individual characteristics and guilt judgments. *Psychology, Public Policy, and Law, 20,* 109-134. doi:10.1037/law0000006

Devine, D. J., Clayton, L. D., Dunford, B. B., Seying, R., & Pryce, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law, 7,* 622-727. doi:10.1037//1076-8971.7.3.622

Dwyer, R. G., & Frierson, R. L. (2006). The presence of low IQ and mental retardation among murder defendants referred for pretrial evaluation. *Journal of Forensic Science, 51,* 678-682. doi:10.1111/j.1556-4029.2006.00115.x

Eisenberg, J. R. (2004). *Law, psychology, and death penalty litigation.* Sarasota, FL: Professional Resource Press.

Elwork, A., Sales, B. D., & Alfini, J. J. (1982). *Making jury instructions understandable.* Charlottesville, VA: Michie.

Elwork, A., Alfini, J. J., & Sales, B. D. (1987). Toward understandable jury instructions. *In the Jury Box: Controversies in the Courtroom,* 161-179.

Elwork, A., Sales, B. D., & Alfini, J. J. (1977). Juridic decisions: In ignorance of the law or in light of it? *Law and Human Behavior, 1,* 163-189. doi:10.1007/BF01053437

Ex parte Briseno, 135 S.W.3d 1, 13 (Tex. Crim. App.).

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods, 39,* 175-191.

Feller, T. (2010). What the literature tells us about the jury foreperson. *The Jury Expert, 22,* 42-54. Retrieved from http://www.thejuryexpert.com/wp-content/uploads/FellerNov2010Vol22Num6.pdf

Fitzgerald, R., & Ellsworth, P. C. (1984). Due process vs. crime control: Death qualification and jury attitudes. *Law and Human Behavior, 8,* 31-51.

Forston, R. F. (1975). Sense and non-sense: Jury trial communication. *Brigham Young University Law Review, 601,* 601-637.

Furman v. Georgia, 408 U.S. 238 (1972). Retrieved from http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=408&invol=238

Garcia, J. G., de la Fuente Solana, E. I., Martinez Arias, M. D. R., & de la Fuente Sanchez, L. (2004). Cross-validation study with confirmatory factorial analysis of the revised legal attitudes questionnaire. *Quality & Quantity, 38,* 391-406.

Garvey, S. P. (1998). Aggravation and mitigation in capital cases: What do jurors think? *Columbia Law Review, 98,* 1538-1576.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th Ed.). Boston: Allyn & Bacon.

Gibbons, F. X., Gibbons, B. N., & Kassin, S. M. (1981). Reactions to the criminal behavior of mentally retarded and nonretarded offenders. *American Journal of Mental Deficiency, 3,* 235-242.

Gillespie, L. K., Smith, M. D., Bjerregaard, B., & Fogel, S. J. (2014). Examining the impact of proximate culpability mitigation in capital punishment sentencing recommendations: The influence of mental health mitigators. *American Journal of Criminal Justice, 39,* 698-715. doi:10.1007/s12103-014-9255-5

Goodman-Delahunty, J., Greene, E., & Hsiao, W. (1998). Construing motive in videotaped killings: The role of jurors' attitudes toward the death penalty. *Law and Human Behavior, 22,* 257-271.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26,* 213-224.

Graham, S., Weiner, B., & Zucker, G. S. (1997). An attributional analysis of punishment goals and public reactions to O.J. Simpson. *Personality and Social Psychology Bulletin, 23,* 331-346.

Greenspan, S., & Switzky, H. N. (2003). Execution exemption should be based on actual vulnerability, not disability label. *Ethics & Behavior, 13,* 19-36.

Gregg v. Georgia, 428 U.S. 153 (1976). Retrieved from http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=428&invol=153

Hall v. Florida, No. 12-10882 (Orally argued March 3, 2014). Retrieved from http://www.law.cornell.edu/supct/cert/12-10882

Haney, C. (2008). Evolving standards of decency: Advancing the nature and logic of capital mitigation. *Hofstra Law Review, 36,* 835-842.

Haney, C., Sontag, L., & Costanzo, S. (2010). Deciding to take a life: Capital juries, sentencing instructions, and the jurisprudence of death. *Journal of Social Issues, 50,* 149-176. http://dx.doi.org.libdata.lib.ua.edu/10.1111/j.1540-4560.1994.tb02414.x

Hauser, D. J., & Schwarz, N. (2015). It's a trap! Instructional manipulation checks prompt systematic thinking on "tricky" tasks. *SAGE Open, April-June,* 1-6. doi:10.1177/2158244015584617

Herr, N. R. (2006, February). Mediation with dichotomous outcomes. Retrieved from http://www.nrhpsych.com/mediation/logmed.html

Heuer, L., & Penrod, S. (1989). Instructing jurors: A field experiment with written and preliminary instructions. *Law and Human Behavior, 13,* 409-430.

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin, 140,* 339-373. doi:10.1037/a0034221

Jacobs, D., & Carmichael, J. T. (2004). Ideology, social threat, and the death sentence: Capital sentences across time and space. *Social Forces, 83,* 249-278. doi:10.1353/sof.2004.0115

Judicial Council of California, Administrative Office of the Courts by the National Center for State Courts (2004). *Does Jury Size Matter? A Review of the Literature.* Retrieved from http://www.ncsc-jurystudies.org/~/media/Microsites/Files/CJS/What%20We%20Do/DOES%20JURY%20SIZE%20MATTER.ashx

Jurek v. Texas, 428 U.S. 262 (1976). Retrieved from https://supreme.justia.com/cases/federal/us/428/262/case.html

Kan, L., Boccaccini, M. T., McGorty, A., Noland, R. M., & Lawson, K. (2009). Presenting information about mental retardation in the courtroom: A content analysis of pre-Atkins capital trial transcripts from Texas. *Law & Psychology Review, 33,* 1-28.

Kane, H. (2003). Straight talk about IQ and the death penalty. *Ethics & Behavior, 13,* 27-33.

Keene, D. (2010, March). Simply jury persuasion: Make your expert optimally persuasive. *The Jury Room.* Retrieved 25 October 2015 from http://keenetrial.com/blog/2010/03/12/simple-jury-persuasion-make-your-expert-optimally-persuasive/

Kenny, D. A. (2014, October 29). Mediation. Retrieved from http://davidakenny.net/cm/mediate.htm

Kerr, N. L., & MacCoun, R. J. (1985). The effects of jury size and polling method on the process and product of jury deliberation. *Journal of Personality and Social Psychology, 48,* 349-363.

Kravitz, D. A., Cutler, B. L., & Brock, P. (1993). Reliability and validity of the original and revised legal attitudes questionnaire. *Law and Human Behavior, 17,* 661-677.

Kroh, M. (2007). Measuring left-right political orientation: The choice of response format. *Public Opinion Quarterly, 71,* 204-220. doi:10.1093/poq/nfm009

Larsen, L., Hartmann, P., & Nyborg, H. (2008). The stability of general intelligence from early adulthood to middle-age. *Intelligence, 36,* 29-34. doi:10.1016/j.intell.2007.01.001

Lieberman, J. D., & Arndt, J. (2000). Understanding the limits of limiting instructions: Social psychological explanations for the failures of instructions to disregard pretrial publicity and other inadmissible evidence. *Psychology, Public Policy, and Law, 6,* 677-711. doi:10.1037//1076-8971.6.3.677

Lockett v. Ohio, 438 U.S. 586, 98 S. Ct. 2954 (1978). Retrieved from https://supreme.justia.com/cases/federal/us/438/586/case.html

Luginbuhl, J., & Middendorf, K. (1988). Death penalty beliefs and jurors' responses to aggravating and mitigating circumstances in capital trials. *Law and Human Behavior, 12,* 263-281.

Lynch, M., & Haney, C. (2009). Capital jury deliberation: Effects on death sentencing, comprehension, and discimination. *Law and Human Behavior, 33,* 481-496. doi:10.1007/s10979-008-9168-2

MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review, 17,* 144-158.

McCann, S. J. (2008). Societal threat, authoritarianism, conservatism, and U.S. state death penalty sentencing (1977-2004). *Journal of Personality and Social Psychology, 94,* 913-923. doi:10.1037/0022-3514.94.5.913

McKoy v. North Carolina. 110 S. Ct. 1227, 108 L.Ed.2d 369 (1990). Retrieved from http://scholarlycommons.law.wlu.edu/cgi/viewcontent.cgi?article=1059&context=wlucdj

Mize, G. E., Hannaford-Agor, P., & Waters, N. L. (2007). *The state-of-the-states survey of jury improvement efforts: A compendium report.* Williamsburg, VA: National Center for State Courts. Retrieved from http://www.ncsc-jurystudies.org/~/media/Microsites/Files/CJS/SOS/SOSCompendiumFinal.ashx

Murray, J., & Farrington, D. P. (2010). Risk factors for conduct disorder and delinquency: Key findings from longitudinal studies. *The Canadian Journal of Psychiatry, 55,* 633-642.

Nagle, J. E., Brodsky, S. L., & Weeter, K. (2014). Gender, smiling, and witness credibility in actual trials. *Behavioral Sciences and the Law, 32,* 195-206. doi:10.1002/bsl.2112

Najdowski, C. J., Bottoms, B. L., & Vargas, M. C. (2009). Jurors' perceptions of juvenile defendants: The influence of intellectual disability, abuse history, and confession evidence. *Behavioral Sciences and the Law, 27,* 401-430. doi:10.1002/bsl.873

Narby, D. J., Cutler, B. L., & Moran, G. (1993). A meta-analysis of the association between authoritarianism and jurors' perceptions of defendant culpability. *Journal of Applied Psychology, 78,* 34-42.

O'Neil, K. M., Patry, M. W., & Penrod, S. D. (2004). Exploring the effects of attitudes toward the death penalty on capital sentencing verdicts. *Psychology, Public Policy, and Law, 10,* 443-470. doi:10.1037/1076-8971.10.4.443

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45,* 867-872. doi:10.1016/j.jesp.2009.03.009

Panek, P. E., & Jungers, M. K. (2008). Effects of age, gender, and causality on perceptions of persons with mental retardation. *Research in Developmental Disabilities, 29,* 125-132. doi:10.1016/j.ridd.2007.01.002

Pellegrino, R. J., Queirolo, S. S., Monarrez, V. E., and Valenzuela, D. M. (1997). Political identification and perceptions of homelessness: Attributed causality and attitudes on public policy. *Psychological Reports, 80,* 1139-1148.

Poulson, R. L., Wuensch, K. L., Brown, M. B., & Braithwaite, R. L. (1997). Mock jurors' evaluations of insanity defense verdict selection: The role of death penalty attitudes. *Journal of Social Behavior and Personality, 12,* 1065-1078.

Prager, I. G., Deckelbaum, G., & Cutler, B. L. (1989). Improving juror understanding for intervening causation instructions. Forensic Reports, 2(3), 187-193.

Proffitt v. Florida, 428 U.S. 242 (1976). Retrieved from https://supreme.justia.com/cases/federal/us/428/242/

Ring v. Arizona, 536 U.S. 584 (2002). Retrived from https://www.law.cornell.edu/supct/html/01-488.ZS.html

Robinson, R. (1995). Irreconcilable differences: Yet more attitudinal discrepancies between death penalty opponents and proponents: A California sample. *Pepperdine Law Review, 22(4),* 1365-1372.

Roper v. Simmons, 543 U.S. 551 (2005). Retrieved from http://www.law.cornell.edu/supct/html/03-633.ZS.html

Saks, M. J., & Marti, M. W. (1997). A meta-analysis of the effects of jury size. *Law and Human Behavior, 21,* 451-467.

Sandys, M., Pruss, H. C., & Walsh, S. M. (2009). Aggravation and mitigation: Findings and implications. *The Journal of Psychiatry & Law, 37,* 189-235.

Schroeder, J., Guin, C. C., Pogue, R., & Bordelon, D. (2006). Mitigating circumstances in death penalty decisions: Using evidence-based research to inform social work practice in capital trials. *Social Work, 51,* 355-364.

Scior, K., Addai-Davies, J., Kenyon, M. & Sheridan, J.C. (2013). Stigma, public awareness about intellectual disability and attitudes to inclusion among different ethnic groups. *Journal of Intellectual Disability Research, 57,* 1014-1026.

Scior, K. (2011). Public awareness, attitudes and beliefs regarding intellectual disability: A systematic review. *Research in Developmental Disabilities, 32,* 2164-2182. doi:10.1016/j.ridd.2011.07.005

Sears, P. M., Pomerantz, A. M., Segreist, D. J., & Rose, P. (2011). Beliefs about the biological (vs. nonbiological) origins of mental illness and the stigmatization of people with mental illness. *American Journal of Psychiatric Rehabilitation, 14,* 109-119. doi:10.1080/15487768.2011.569665

Severance, L. J., & Loftus, E. F. (1982). Improving the ability of jurors to comprehend and apply criminal jury instructions. *Law & Society Review, 17,* 153-197. doi:10.2037/3053535

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

Severance, L. J., Greene, E., & Loftus, E. F. (1984). Toward criminal jury instructions that jurors can understand. *Journal of Criminal Law & Criminology, 75,* 198-233. doi:10.2307/1143210

Shaver, K. G. (1985). The attribution of blame: Causality, responsibility, and blameworthiness. New York: Springer-Verlag.

Silverman, W., Miezejeski, C., Ryan, R., Zigman, W., Krinsky-McHale, S., & Urv, T. (2010). Stanford-Binet and WAIS IQ differences and their implications for adults with intellectual disability (aka mental retardation). *Intelligence, 38,* 242-248. doi:10.1016/j.intell.2009.12.005

Simon, C. L., & Clopton, J. R. (1984). Comparisn of WAIS and WAIS-R scores of mildly and moderately mentally retarded adults. *American Journal of Mental Deficiency, 89,* 301-303.

Slater, P. (1958). Contrasting correlates of group size. *Sociometry, 21,* 129-139.

Smith, A. E., & Haney, C. (2011). Getting to the point: Attempting to improve juror comprehension of capital penalty phase instructions. *Law and Human Behavior, 35,* 339-350. doi:10.1007/s10979-010-9246-0

Spitz, H. H. (1986). Disparities in mentally retarded persons' IQs derived from different intelligence tests. *American Journal of Mental Deficiency, 5,* 588-591.

Stevens, K. B., & Price, J. R. (2006). Adaptive behavior, mental retardation, and the death penalty. *Journal of Forensic Psychology Practice, 6,* 1-29. doi:10.1300/J158v06n03_01

Stevenson, M. C., Bottoms, B. L., & Diamond, S. S. (2010). Jurors' discussions of a defendant's history of child abuse and alcohol abuse in capital sentencing deliberations. *Psychology, Public Policy, and Law, 16*(1), 1-38. doi:10.1037/a0018404

Tassé, M. J. (2009). Adaptive behavior assessment and the diagnosis of mental retardation in capital cases. *Applied Neuropsychology, 16,* 114-123. doi:10.1080/09084280902864451

Tetterton, V. S., & Brodsky, S. L. (2007). More is sometimes better: Increased mitigating evidence and sentencing leniency. *Journal of Forensic Psychology Practice, 7,* 79-85. doi:10.1300/J158v07n03_05

Thompson, W. C., Cowan, C. L., Ellsworth, P. C., & Harrington, J. C. (1984). Death penalty attitudes and conviction proneness: The translation of attitudes into verdicts. *Law and Human Behavior, 8,* 95-113. doi:10.1007/BF01044353

Thompson v. Oklahoma, 487 U.S. 815 (1988). Retrieved from http://www.law.cornell.edu/supremecourt/text/487/815

Trahan, A. (2011). In their own words: Capital jurors' reactions to mitigation strategies. *International Journal of Punishment & Sentencing, 7,* 1-16.

U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention (2004). Risk factors for delinquency: An overview. (NJC Number: NJC 207540). Retrieved from: https://www.ncjrs.gov/pdffiles1/ojjdp/frd030127.pdf

Vogel, B. L. (2003). Support for life in prison without the possibility of parole among death penalty proponents. *American Journal of Criminal Justice, 27,* 263-275.

Wainwright v. Witt, 469 U.S. 412 (1985). Retrieved from http://supreme.justia.com/cases/federal/us/469/412/

Weber, R., & Popova, L. (2012). Testing equivalence in communication research: Theory and application. *Communication Methods and Measures, 6,* 190-213. doi:10.1080/19312458.2012.703834

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92,* 548-573.

Weiner, B. (2006). *Social motivation, justice, and the moral emotions: an attributional approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

Weiner, B., Osborne, D., & Rudolph, U. (2010). An attributional analysis of reactions to poverty: The political ideology of the giver and the perceived morality of the receiver. *Personality and Social Psychology Review, 15,* 199-213. doi:10.1177/1088868310387615

Wellek, S. (2010). *Testing statistical hypotheses of equivalence and noninferiority, second edition.* Chapman & Hall/CRC: Boca Raton, FL.

Whitaker, S. (2008). The stability of IQ in people with low intellectual ability: An analysis of the literature. *Intellectual and Developmental Disabilities, 46,* 120-128.

White, L. T. (1987). Juror decision making in the capital penalty trial: An analysis of crimes and defense strategies. *Law and Human Behavior, 11,* 113-130.

Williams v. Florida, 399 U.S. 78 (1970). Retrieved from https://supreme.justia.com/cases/federal/us/399/78/case.html

Williams v. Quarterman, No. 05-20350 (5th Cir. December 09, 2008). Retrieved from http://caselaw.findlaw.com/us-5th-circuit/1239187.html

Witherspoon v. Illinois, 391 U.S. 510 (1968). Retrieved from https://supreme.justia.com/cases/federal/us/391/510/case.html

Wolbransky, M. F. (2011). *Capital jury sentencing recommendations: The relationship between mental illness-related mitigating factors and life versus death decisions (*Unpublished doctoral dissertation). Drexel University, Philadelphia, PA.

Yarnell, M. A. (2005). The Arizona jury past, present and future reform. Retrieved from http://www.michaelyarnell.com/TheArizonaJury-Full.pdf

Ziller, R. C. (1957). Group size: A determinant of the quality and stability of group decisions. *Sociometry*, 165-173.

Zucker, G. S., & Weiner, B. (1993). Conservatism and perceptions of poverty: An attributional analysis. *Journal of Applied Social Psychology, 23*, 925-943. doi:10.1111/j.1559-1816.1993.tb01014.x

APPENDICES

APPENDIX A

Jury Instructions

Mr. Michael Stevens has been found guilty of murder during a first degree robbery. You will decide the punishment for Mr. Stevens – either life in prison without the possibility of parole (LWOP) or death. The punishment you choose will depend on whether the aggravating factor outweighs the mitigating factors.

An aggravating factor supports sentence of death while a mitigating factor supports a sentence of LWOP. You will weigh the aggravating and mitigating factors to determine his punishment. The judge must consider your sentence in deciding Mr. Stevens's punishment.

The jury who found Mr. Stevens guilty decided the following aggravating factor – that the crime was committed during a robbery. You should consider this in deciding his sentence.

Mr. Stevens will offer evidence of mitigating factors that support a sentence of LWOP. There is no burden of proof for these factors.

Mitigating factors can include any aspect of Mr. Stevens's character or background including abuse or neglect, adaptive behavior, and intelligence. You should only rely on the evidence and the law - not on passion, prejudice, or any other arbitrary factors.

Your job is to weigh the mitigating against the aggravating factors to determine the sentence for Mr. Stevens. You may weigh some factors more than others.

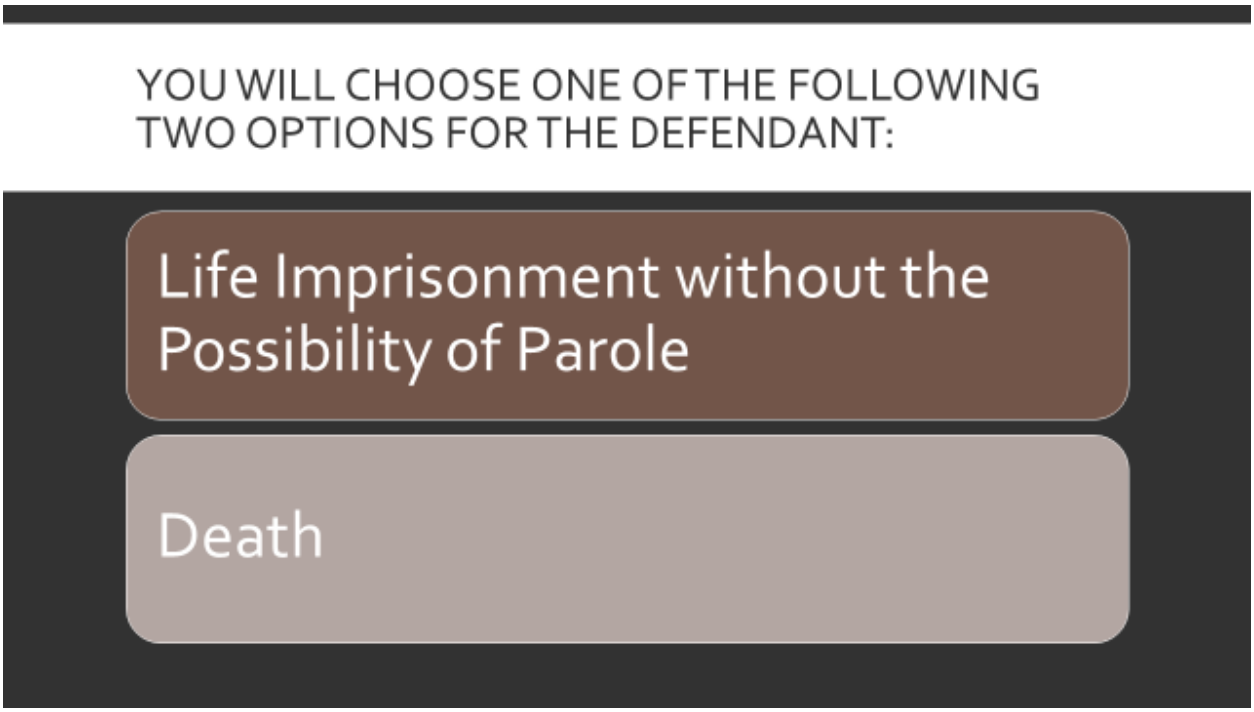
All of you must agree whether Mr. Stevens should be sentenced to LWOP or death. The judge is required to consider a sentence of LWOP as a mitigating factor.

If you are convinced that the aggravating factor outweighs the mitigating factors, your verdict would be a sentence of death.

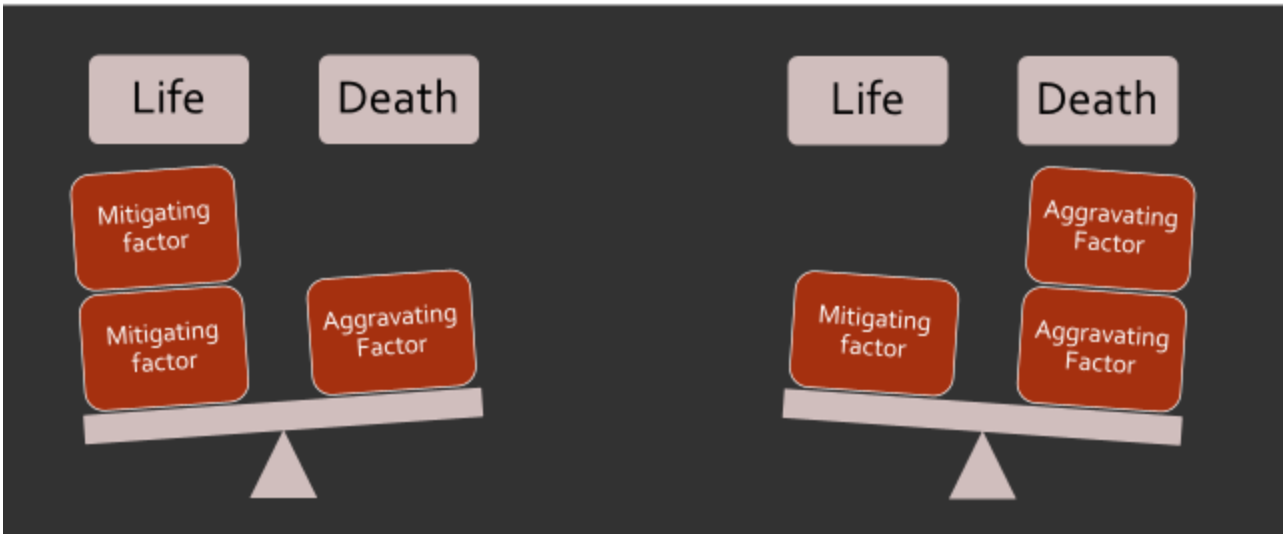
However, if you are not convinced that the aggravating factor outweighs the mitigating factor, your verdict would a sentence of LWOP.

APPENDIX B

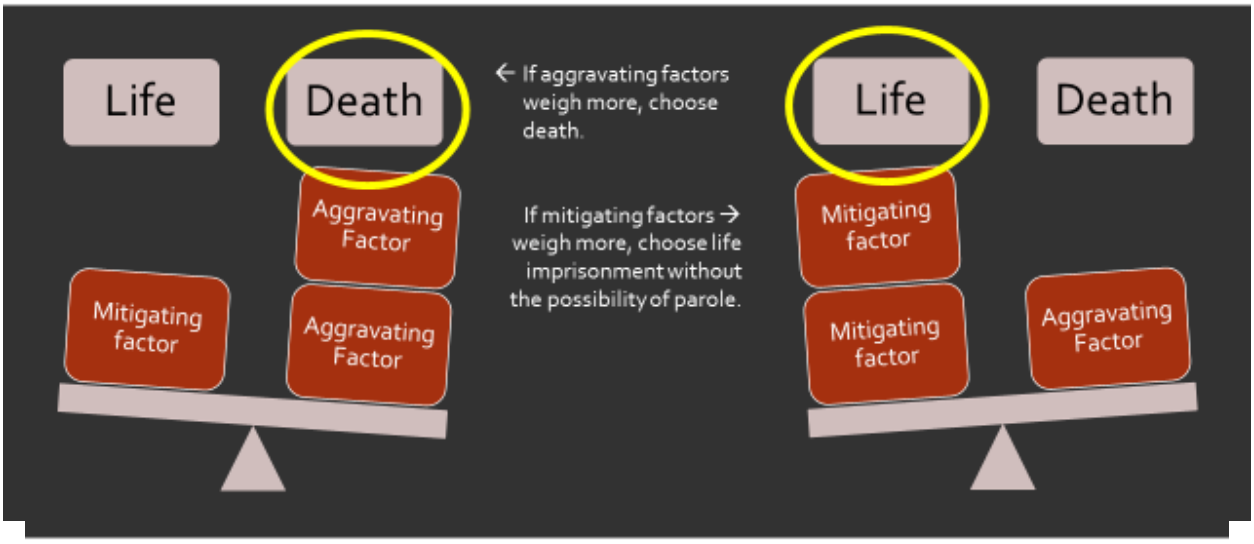
PowerPoint Supplement to Jury Instructions



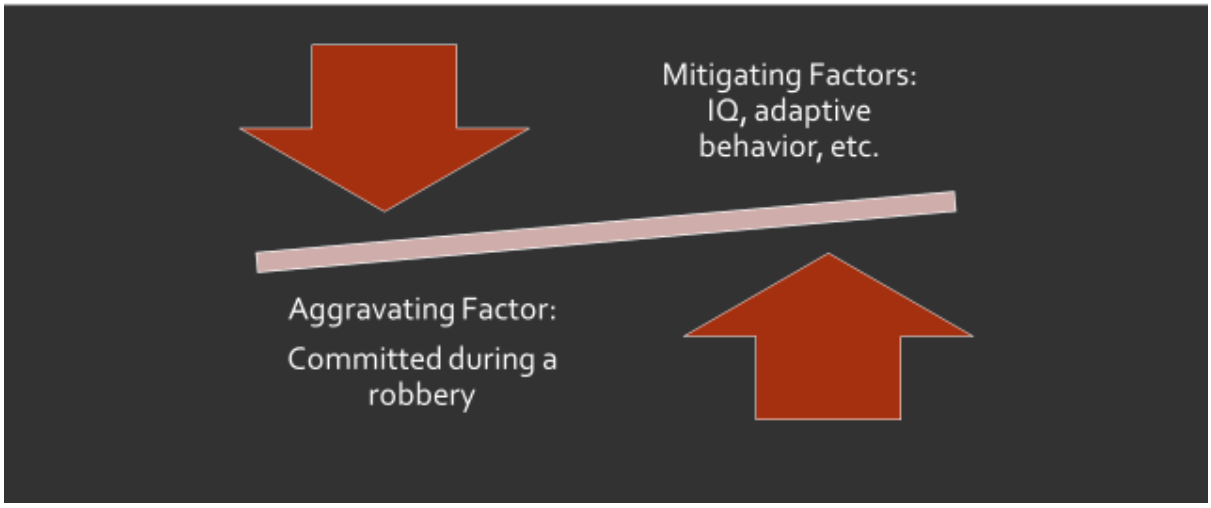
WEIGH THE AGGRAVATING & MITIGATING FACTORS



YOU ALL MUST AGREE – UNANIMITY IS REQUIRED.



WEIGHING THE FACTORS



APPENDIX C

Demographic Form

Participant ID: ____

Jury Number: ____

Please fill out the following questions to the best of your ability. Please answer every question.

Age (in years): ____

Gender (circle one): Male Female

Race/Ethnicity: _____

Place of Birth (City, State): _____

Place Where Raised* (City, State): _____

**In other words, where would you say you spent most of your childhood?*

Political Orientation:

1 2 3 4 5 6 7 8 9 10 11

Left/Liberal Orientation *Right/Conservative Orientation*

Political Affiliation:

Democrat Republican None Other (specify): _____

Number of Years of College <u>Completed</u>: _____

Academic Major: _____

College GPA: _____

High School GPA: _____

Have you ever served on a jury before? Yes No

If yes, please provide details about that experience including the type of trial (i.e., civil vs. criminal), the type of crime, the location, the verdict, etc.

_____

_____

_____

Have you ever been the victim of a violent crime before? Yes No

Has anyone close to you ever been the victim of a violent crime? Yes No

If you feel comfortable doing so, please provide details about the experience noted above including your relationship to the victim, the type of violent crime, the outcome (i.e., were charges filed? Was the perpetrator convicted?), etc.

_____

_____

_____

APPENDIX D

Transcript without Manipulations

RE-SENTENCING HEARING IN THE CASE OF STATE OF ALABAMA VERSUS MICHAEL STEVENS FOR THE CHARGE OF CAPITAL MURDER (MURDER DURING THE COMMISSION OF A ROBBERY): Testimony by a forensic clinical psychologist hired by the defense

Q. Dr. Thompson, what were you asked to do in this case?

A. I was asked to conduct a mitigation evaluation.

Q. Could you please tell the court what that means and what you did in this case?

A. A mitigation evaluation involves the identification of factors that might suggest a defendant is more or less deserving of a punishment of death or life in prison. For this evaluation, I interviewed the defendant, his family, and his friends. I also looked at records and administered some tests.

Q. Now, you have testified that part of a forensic assessment involves looking at things that might be raised in mitigation; is that right?

A. Yes.

Q. When you're looking for such factors, what are you looking for?

A. What I'm looking for are both risk and protective factors in an individual's life.

Q. What are risk factors?

A. Risk factors increase the chance that someone will offend later in life. In all cases, I look at the person's history and try to understand how they became a person who could be charged with capital murder. There are many risk factors that could have negatively influenced a person's development. I ask about things like parental discipline, emotional attachment to parents, parental investment, school experiences, the types of people that they hung around with, access to weapons or drugs, and a variety of other things that influence how we become the people that we are. We are all raised over the course of time with good and bad influences; I just try to get a sense of how this individual, Mr. Stevens, was raised.

Q. What are protective factors?

A. Protective factors are those factors that could positively influence a person's development. In terms of protective factors, I'm looking for positive things that could have assisted, or did assist,

Mr. Stevens in coping in various situations. I'm looking at parental role models, parental discipline, community involvement, environment, and friendships. Protective factors may reduce the negative impact of risk factors.

Q. Does the research show anything regarding the impact of risk factors on future behavior and eventual success?

A. The research shows the more risk factors to which someone is exposed and the earlier they experience them, the worse the prognosis. Many risk factors lead to dysfunction, unless enough protective factors are in place to assist in coping.

Q. And can early intervention have an impact on the risk factors, if present? And if so, how?

A. Early intervention can assist an individual in developing coping strategies. For example, coping with and intervening with problems to get beyond the difficulties and not succumb to negative outcomes. There is a lot of research that says that early intervention is effective in reducing future problem behaviors, including criminal and violent behavior, but also in developing coping strategies that are adaptive for the individual.

Q. Did you find risk and protective factors that applied to Mr. Stevens during development?

A. Yes, I did.

Q. Doctor, is this a relatively comprehensive list of risk factors present in Mr. Stevens' life?

A. Yes.

Q. Okay, will you please go through each one, explain what it is, and let the court know how it relates to Mr. Stevens' upbringing and how you believe it impacted him?

****INSERT MANIPULATIONS HERE****

Q. Okay, moving on to childhood onset psychological disorder. What is ADHD?

A. ADHD is s a common condition that affects children and adolescents and can continue into adulthood for some. Children with ADHD generally have problems paying attention or concentrating. They can't seem to follow directions and are easily bored or frustrated with tasks. They also tend to move constantly and are impulsive, not stopping to think before they act. These behaviors are generally common in children. But they occur more often than usual and are more severe in a child with ADHD. The behaviors that are common with ADHD interfere with a child's ability to function at school and at home. Adults with ADHD may have difficulty with time management, organizational skills, goal setting, and employment. They may also have problems with relationships, self-esteem, and addictions.

Q. Tell the Court what you know about the effects of ADHD on Mr. Stevens.

A. Well, many kids with ADHD have problems with learning, just like Mr. Stevens did, which is made worse by their problems with organization and forgetfulness. It is not a learning disability, per se, but some studies have shown that as many as half of the kids with ADHD also have a learning disability. Kids with ADHD also get in trouble more frequently than their peers because they tend to demonstrate hyperactive and disruptive behaviors in the classroom. It can also be difficult to parent a child with ADHD because of the level of impulsivity they often demonstrate. This impulsivity and hyperactivity also makes it difficult to establish and maintain relationships with peers. In short, kids with ADHD tend to struggle more in academic, home, and social settings relative to their peers. This is consistent with Mr. Stevens who had troubles in all of these domains.

Q. You said that he received treatment for the disorder. Tell the Court about the treatment that he received.

A. When he was 7 years old he briefly participated in behavior modification therapy. Mr. Stevens was uncooperative with treatment and often refused to go. He would not participate, despite encouragement from his therapist and his mother. The therapy was terminated shortly after it started because of his poor cooperation. In short, he would have benefited from more frequent and intensive treatment. So although he was in treatment, it was not effective.

Q. Why is that a risk factor?

A. Well, it is not uncommon for children with ADHD, those who are not adequately treated, to increase their problem behaviors in school and at home. Sometimes they are suspended from school which leads to numerous problems including poor academic performance, a tendency to hang around with children who get into similar troubles, and strained relationships with parents. Sometimes the level of parent-child conflict gets to the point where parents give up trying to control their children's behavior. It appears that this was this case for Mr. Stevens. He was so uncooperative and noncompliant that his mother eventually stopped trying to keep him in treatment.

Q. Tell me about Mr. Stevens's home life.

A. Well, the home was in a lower middle class neighborhood and it was small, but it was kept tidy and clean. Mr. Stevens's mother worked as a nurse at the hospital and took care of Mr. Stevens and his older brother by herself. Although the boys saw their father on occasional

weekends and holidays, he was rarely involved other than to pay child support. He had his own family and other children to take care of so Mr. Stevens and his brother did not see much of him. Mr. Stevens's mother did her best to take care of her children and she provided as much support to them as she could even though she was a single parent who was working full-time as a nurse. She tried to get her sons involved in sports and after school activities. Mr. Stevens's older brother, Julius, was involved in a number of activities and eventually went on to play football at a community college. Mr. Stevens, on the other hand, had little interest in being involved in these activities and spent most of his time with friends he made at school who were from the other side of town. As a result, Mr. Stevens did not spend much of his time at home. His mom would often drive to these other areas to pick up her son, but he would refuse to come home with her. He found his niche with these other kids.

Q. Tell me about these friends he had and the places where he would hang out with his friends.

A. Well, the friends he made were from a lower socioeconomic status than he and his family. In these neighborhoods, Mr. Stevens was exposed to things he was not exposed to at home like gang activity, substance abuse, and criminal activity. He eventually joined a gang at age 14 and this is where he received most of his emotional and social support. He spent most of his days in these neighborhoods because he was getting suspended or skipping school altogether. Mr. Stevens was often unsupervised during these times, or, if there was supervision, it was by the parents of his friends who were also involved in criminal activity and using drugs. His mother and older brother did their best to encourage Mr. Stevens to stay home and spend time with them, but it seemed as though Mr. Stevens felt like he connected with these people and that he could rely on them and trust them to support him. He would leave home and refuse to come home for days at a time.

Q. Doctor, why is this relevant to the development of Mr. Stevens?

A. Well, we learn a lot from our families and communities. The places where Mr. Stevens chose to spend most of his time modeled negative behaviors such as criminal activity, substance abuse, and violence. He learned from hanging around these areas that crime was okay and handling problems with violence was appropriate. He saw others get reinforced for these behaviors and was not exposed to the typical consequences that often follow criminal behavior. Mr. Stevens was carrying a gun for protection by the age of 15. The research has shown that exposure to

things like crime and violence is predictive of other behaviors such as substance abuse, adolescent criminal behavior, and eventual perpetuation of violence.

Q. You just mentioned substance abuse – tell the Court more about this and how it is relevant in this case.

A. Mr. Stevens started to use drugs and alcohol at a young age, about 14 or 15. Drugs and alcohol were readily available to him through friends and he started to sell drugs around the same time to make some extra money. In fact, almost all of his peers used drugs and alcohol and so did their parents. Again, we see he was modeling the negative behaviors of those around him. Although there was never alcohol or drugs in his own home, he often saw adults in his friends' neighborhoods engage in alcohol and drug use regularly and often to excess.

Q. The next risk factor on the list is parenting. Tell us about this.

A. Mrs. Stevens worked full-time as a nurse at the hospital and sometimes had to work the late shift of pick up extra hours. Because she was working these extra hours, she was not able to spend as much time with her children as she would have liked. She did what she thought was best in terms of providing for her family and did well considering she was the only parent raising those kids. She hired babysitters when the kids were younger, but by the time her eldest turned 12, he was responsible for watching Mr. Stevens. Of course, he could not do this very well and, not surprisingly, Mr. Stevens would get involved in mischief when he was in the care of his older brother. By the time he was a teenager, Mr. Stevens's simple mischief, which is typical of young kids, turned into crime. When Mrs. Stevens was able to be home with the kids, she was often tired and worn out from her long hours.

Q. Did you see any indication as to whether Mrs. Stevens was aware of what was going on in her son's lives?

A. According to records, she did her best to be involved but seemed to fall short. In addition, Mr. Stevens was not receptive to her involvement in his life so he would push her away. When I spoke with Mr. Stevens's mom, she indicated that she wanted to be around but simply could not be because she had a family to provide for. She also stated that she feels responsible for what has happened and feels great sadness about her absence during her son's upbringing. All these things I have mentioned are risk factors for conduct disorder.

Q. What is conduct disorder?

A. Conduct disorder is a behavioral disorder in adolescents. Kids with conduct disorder display long-standing behavior problems including violating others' rights and the norms of society. Symptoms include behavior problems including violence and destruction of property. Sometimes kids with conduct disorder lie and break rules as well.

Q. And why is that relevant in this case?

A. Without intervention for things like limited parenting, substance abuse, and exposure to crime and violence, the likelihood of developing conduct disorder is high. These factors also influence functioning in school and with friends. The end result for someone with this many risk factors is often the criminal justice system. Because Mrs. Stevens was busy with work, she was not able to spend much time with her kids and when she was around, Mr. Stevens was resistant to her involvement because he wanted to spend time with his friends on the other side of town.

Q. The next risk factor is father absence. What does that refer to?

A. As I mentioned before, Mr. Stevens's biological father was rarely around because he had a wife and kids of his own. He provided financial support and visited the boys on holidays and some weekends when he could. Mr. Stevens frequently asked about his father and wanted a relationship with him, but it was not feasible because he lived out of town and had another family to take care of. I think it was hard for Mr. Stevens to understand that his father had other responsibilities.

Q. What are the implications of his absence?

A. Boys tend to look up to their dads and idolize them to some degree. They want to be like their fathers. Mr. Stevens wanted to know his father and he didn't have that chance. There was not a positive male role model in his life to fulfill the role of a father figure to him. He did have plenty of examples of negative male influences from the interactions with his friends and other gang members. Most of the other people in the gang were much older than him; in some cases, they were 15 and 20 years older and modeled criminal behavior and violence. Mr. Stevens idolized these guys he hung around, it was all he knew. He looked up to these guys and did not have any examples of positive male role models or positive coping. Even though his older brother could have been a positive role model in his life, Mr. Stevens spent the majority of his time with these other kids outside of the home.

Q. Why is father absence a risk factor?

A. It's a risk factor for boys particularly, but really for any child. Research has shown that boys who grow up without a father figure at home have more trouble in school, end up in more trouble with the legal system, and are at an increased risk for mental illness and conduct disorder. Absence of a male role model negatively impacts the development of young men.

Q. Okay, moving on to the next two risk factors. Tell the Court how a failure to provide supervision and guidance impacts a person's development.

A. His older brother, one of the only positive role models in his life, was often at football practice at the school or tutoring sessions that were provided by the school for student athletes. There was not much guidance either. Ms. Stevens was gone most of the time and when she was home, she was tired. I mentioned this already, but she feels guilty about not being there for her children. She did not provide them with guidance. They were left to find activities on their own, to meet friends and adult influences by themselves. This sort of neglect is a very powerful risk factor, particularly when combined with other risk factors. Neglect consistently predicts delinquency.

Q. What are the implications from child neglect? Why is that a risk factor?

A. Well, neglected children struggle socially. They have trouble making and keeping friends and this is in part due to the lack of modeling of appropriate ways to connect with people. When children's needs are not met they go into other environments and make social connections and in this case he did and these connections were with people who were entrenched in the criminal lifestyle. Neglected children desire attention and tend to find someone, anyone, to bond with – regardless of whether they are positive or negative.

Q. The next issue is youthfulness. What does that refer to?

A. Mr. Stevens was only nineteen at the time of the crime. Research has shown that the human brain is not fully developed by then. The frontal lobe, which is responsible for impulse control and decision-making, is one of the last areas to develop and usually isn't fully developed until the early to mid-20s. As such, Mr. Stevens' ability to function like an adult in these areas was limited due to his age.

The other thing to realize is that individuals with ADHD, conduct disorder, and lower intellectual functioning tend to mature more slowly because these problems don't allow them to mature in the same way as typically-developing adolescents. When you add this to the fact that

his role models were involved in crime and substance use – his youthfulness really limited his ability to make good decisions and to behave appropriately.

Q. Your chart lists risk factors and protective factors. Doctor, please tell the court what the term resilience refers to?

A. Resilience is an individual's ability to overcome the negative impact of risk factors. Resilience comes from protective factors. In short, early intervention and protective factors can lead to the development of resilience.

Q. Now the next slide refers to protective factors. Please describe what protective factors are.

A. Protective factors promote healthy adjustment and protect against risk factors. An example would be a parent who is interested in the child and takes time to help them with their homework, meets with the teachers, or is involved with the child in any positive manner.

Q. Who is Joanne Whitley?

A. Joanne Whitley was an adult role model for Mr. Stevens. She was a friend of the family that Mr. Stevens stayed with sometimes when he wanted a break from his other friends or when he wanted to get away from his family. When he stayed there, he lived by her rules and performed well. He did chores and completed his homework. He responded well to the structure she provided for him. She trusted him enough to ask him to take care of her kids while she was gone and would ask him to help her around the house. Mr. Stevens described her as a mother-type figure.

Q. Can the structure that was provided by Ms. Whitley's home allow individuals such as Mr. Stevens who have difficulties with intellectual functioning function well?

A. Yes; structure is one of the most important things for people with lower IQs. The ability to predict what is going to happen and rules and regulations really help them succeed. To some degree it helped him in functioning, but it was late. It would have been much better had it happened earlier in his life.

****INSERT SECOND MANIPULATIONS HERE****

Q. You have talked about all of these protective factors and risk factors. Is there a cumulative effect when you have multiple risk factors?

A. Yes.

Q. Tell us about that.

A. The presence of one risk factor is a problem – but people can generally overcome one problem, but the more risk factors you have, the worse it gets. For example, think about his academic and school failure. Match that with his conduct disorder issues which affected his relationships with his teachers. The risk factors pile upon one and another and interact with each other. The more you get the worse off it is because people don't find a way out and don't have the general coping skills.

With this many problems in place, it is very hard to intervene. This is why I stress the importance of early intervention and effective early intervention. But in a case like this, it is very hard to get enough protective factors in play to change things. Things just get worse and worse and worse without some form of intervention.

Q. And is all of that compounded by one's intelligence?

A. Yes.

Q: That's all I have, Your Honor.

CROSS EXAMINATION

BY MR. HOWELL (DISTRICT ATTORNEY):

Q. Dr. Thompson, you have described a very interesting childhood here. But from the sounds of it, Mr. Stevens chose a lot of his negative peers and the environments he stayed in, is that right?

A. Well, Mr. Stevens was not getting much attention at home, as I mentioned, and he did not have a positive male role model to guide him in the right direction. Mr. Stevens sought social support from others, and he happened to find it with individuals who were negative influences on him. I supposed you could say he chose it, however, you have to look at what he was working with and the factors that contributed to that decision.

Q. Well you testified that his brother was a star athlete who played football at the community college. Why is it that Mr. Stevens turned out so different if they came from the same home, doctor?

A. I testified about the factors that influenced Mr. Stevens from a young age. He presented as a behavior problem from early elementary school which was related to his diagnosis of ADHD. He had trouble performing well in school and these difficulties impacted him socially and academically. His brother did not have similar problems when he was younger. It was during this time when Mr. Stevens and his brother started to move in different directions. They just preferred different things and were on different paths.

Q. Didn't you say that Mrs. Stevens and his brother wanted Mr. Stevens to come home and stay with them and he would refuse?

A. Yes, he would refuse to come home.

Q. So, Mr. Stevens decided to stay out with his friends even though he had a perfectly good home with a mother who cared for him and worked hard to provide for her family and an older brother who was a positive role model? Am I understanding this correct, doctor?

A. Well, in a sense, yes.

Q. That's what I thought. Was it Mrs. Stevens who reported that Mr. Stevens sold drugs to help his mom pay for things at home?

A. That was Mr. Stevens.

Q. How old was he when this was going on? Did you ask him?

A. It was around 14 or 15 when he got involved with the gang. That was part of his job in the gang and he skimmed money off the top for himself and gave some to his mom to help out.

Q. Did you ask Mr. Stevens about why he committed the crime?

A. I did not get into that with him because it was not a part of my job to look at that. My job was to identify risk factors that played into his development.

Q. So Mr. Stevens did not tell you that he reason he killed the parking attendant was because he wanted a tear drop tattoo and wanted to move up in his gang?

A. No, we did not talk about that. When I asked about his gang activity, I was more interested in learning about those relationships and how he was impacted by them.

Q. Out of curiosity, Doctor, would you say that participation in church activities, Boy Scouts, or vocational training is a protective factor?

A. I would say that, yes. Those can all be positive influences in a child.

Q. Thank you, I think that's all I have.

REDIRECT EXAMINATION BY MR. JONES (DEFENSE ATTORNEY):

Q. Okay. Let me just ask a few clarifying question then. You were not hired to assess whether or not Mr. Stevens is intellectually disabled, correct?

A. I was not.

Q. Do you know whether Mr. Stevens was in the Boy Scouts?

A. My recollection is that he was in the Boy Scouts, but had no longstanding commitment with them. He may have gone once or twice. He never really had any extracurricular activities. Boy Scouts could have been a really good thing, but it turns out it wasn't.

MR. JONES: That's all I have

MR. HOWELL: Nothing further, Your Honor.

APPENDIX E

Condition-Specific Manipulation Testimony Recall

Transcript	Prong 1: Intellectual Functioning	Prong 2: Adaptive Functioning
A	Satisfied	Satisfied
B	Not Satisfied	Not Satisfied
C	Satisfied	Not Satisfied
D	Not Satisfied	Satisfied

Transcript A

Q. Is there anything that you have uncovered that is inconsistent with a diagnosis of intellectual disability?

A. From my review of everything that I have received, there is nothing inconsistent with a diagnosis of intellectual disability.

Q. And to the extent that you saw how Mr. Stevens did or didn't function, nothing you saw is inconsistent with a diagnosis of intellectual disability?

A. Nothing is inconsistent with the diagnosis.

Q. You didn't attempt to determine whether Mr. Stevens was intellectually disabled?

A. No, I did not.

Transcript B

Q. Is there anything that you have uncovered that is inconsistent with a diagnosis of mild intellectual disability?

A. From my review of everything that I have received, Mr. Stevens's IQ score and the fact that his adaptive behavior deficits were not severe enough to meet criteria for deficits were inconsistent with intellectual disability.

Q. And to the extent that you saw how Mr. Stevens did or didn't function, what you saw is inconsistent with a diagnosis of intellectual disability?

A. Yes, these factors are inconsistent with the diagnosis.

Q. You didn't attempt to determine whether Mr. Stevens was intellectually disabled?

A. No, I did not.

Transcript C

Q. Is there anything that you have uncovered that is inconsistent with a diagnosis of mild intellectual disability?

A. From my review of everything that I have received, the fact that his adaptive behavior deficits were not severe enough to meet criteria for deficits were inconsistent with intellectual disability, though his IQ score was consistent with the diagnosis.

Q. And to the extent that you saw how Mr. Stevens did or didn't function, his adaptive behavior was inconsistent with a diagnosis of intellectual disability?

A. Yes, his adaptive functioning was inconsistent with the diagnosis even though his IQ was not.

Q. You didn't attempt to determine whether Mr. Stevens was intellectually disabled?

A. No, I did not.

Transcript D

Q. Is there anything that you have uncovered that is inconsistent with a diagnosis of mild intellectual disability?

A. From my review of everything that I have received, Mr. Stevens's IQ score was inconsistent with intellectual disability, though his adaptive behavior deficits were consistent with the diagnosis.

Q. And to the extent that you saw how Mr. Stevens did or didn't function, Mr. Stevens's IQ score was inconsistent with a diagnosis of intellectual disability?

A. Yes, his IQ score was inconsistent with the diagnosis even though his adaptive behavior deficits were not.

Q. You didn't attempt to determine whether Mr. Stevens was intellectually disabled?

A. No, I did not.

APPENDIX F

IQ Manipulations

IQ DEFICIT CONDITION

Could you tell the court what an intelligence test is and why it is used?

A. An intelligence, or IQ, test is used to determine how someone thinks compared to other people their age. In general, IQ tests measure vocabulary knowledge, reasoning skills, the ability to use information, and how quickly someone thinks. IQ tests are often used in schools to determine if a child needs special services. Students whose IQ is much lower than students who are the same age might qualify for special education, whereas a student whose IQ is much higher might qualify for gifted services. IQ tests are also used to determine if someone is eligible for social security or disability benefits. An individual's score is compared to a larger sample of people the same age. For instance, a 17-year-old would be compared to other 17-year-olds. The average IQ score is 100, meaning that half of the population has an IQ of 100 or below and the other half is above 100. The majority of people in the United States fall within the average range which is from 85 to 115.

Q. Did you review any documents that suggested that Mr. Stevens' had been given an IQ test at some point when he was a child?

A. Yes, Mr. Stevens' IQ was tested when he was 12 years old and in the 5th grade. His verbal IQ score, which reflects his ability to use and understand words and language, was 75. His performance IQ, which is a measure of his visual and spatial skills, was very similar at 73. Visual and spatial skills allow us to look at something and understand it – for example, perceiving images in a mirror or being able to pack a suitcase or complete a puzzle. His overall IQ score, which is a measure of all the skills measured by the test, including the words that he knows, information that is commonly known in the community, and his ability to look at objects and perceive them accurately, was similar to his other two scores – it was 73.

What is important to understand is that all of these scores are well below average – well below what almost everybody in the United States earns. In fact, his scores are low enough that they fall within the range for a diagnosis of intellectual disability – intellectual disability is the new term for the condition which used to be called mental retardation. IQ scores up to 75 are included in the range for intellectual disability. His overall IQ score of 73 also falls in the borderline range which includes scores between 70 and 79. Regardless of which range you call it, only about 2% of the population scores this low on an IQ test.

Q. Doctor, tell us more about what an IQ score of 73 means.

A. Well, most people who take this test score somewhere between 85 and 115, with the majority of people scoring close to 100. When you look at all possible scores, 50% of people score below 100 and 50% above 100. He fell in the bottom 4th percentile meaning that 96% of the children his age scored higher than he did on this test when it was being made.

Q. Does a person's intelligence, as measured by an IQ test, affect the way a person functions at school? Meaning the person's ability to understand topics and participate in classes and interact with students and teachers?

A. Yes, a person's intelligence has a direct impact on a person's ability to learn and to function well in an educational setting. Mr. Stevens' score is very much in-line with his grades. His grades were very low throughout all of his schooling with the majority of his grades being D's, but he also earned some F's in the more difficult courses. His inability to learn and comprehend was to the degree that he failed the 5th grade. What is important to think about is that he was not tested or provided services until the 5th grade. Had he been tested earlier, I am certain that the teachers would have recognized that his intellectual skills were very low and that he needed extra assistance.

Q. Is that sometimes referred to as special education?

A. Yes. In my opinion, because the teachers did not know the extent of his problems in learning they focused on trying to manage his behavior problems. It is not at all surprising that Mr. Stevens didn't behave well in the school setting. He was not doing well in his classes and it was clear that no matter how hard he tried, he couldn't keep up with his friends. Kids like this often give up on learning and spend their time disrupting the classroom. In his case, Mr. Stevens was getting in trouble in school for arguing with teachers, arguing with his peers, and using profanity in the classroom. Typically if kids do well and succeed in the classroom, their behavior improves. If they don't succeed, they tend to act out and get attention in a negative way. According to Mr. Stevens' mother and his 4th grade teacher, her son was socially promoted to the 7th grade and eventually quit school because his grades were so low.

Q. Did you give Mr. Stevens' an IQ test for the current evaluation?

A. I did. Mr. Stevens obtained a full scale IQ of 71which is in the 3rd percentile. That means that only 3% of individuals his age scored lower than he did at the time the test was made. In other words, 97% of his peers scored higher than he did on this test. Similar to prior testing, his full scale IQ falls within the range necessary for a diagnosis of intellectual disability.

Q. Tell us why low intelligence is important to consider when you think about how he developed into the person that he is today, as well as the person that he was at the time of the crime.

A. Research has shown that doing very poorly in school, and doing so from an early age, is a risk factor for many things – such as substance abuse, adolescent criminal behavior, and school

dropout. Basically, children who are not as "smart," if you will, have fewer opportunities in life, often have problems socializing with others, and often feel bad about themselves and these decrease the chances of success during teenage and adult years. With regard to opportunities, individuals like Mr. Stevens, do not typically succeed at sports or clubs such as chess or debate – participation in extracurricular activities such as sports and clubs are protective factors. Instead, people like Mr. Stevens will often stay to themselves, but some will join in with negative peer groups.

If you think about an IQ score as a measure of ability – which psychologists do – you have to consider that Mr. Stevens began school at a disadvantage. He was not as bright as other students and he had more difficulty solving problems and he didn't have the ability to weigh the pros and cons of his behavior in the same way that people with higher intelligence can.

NO IQ DEFICIT CONDITION

Could you tell the court what an intelligence test is and why it is used?

A. An intelligence, or IQ, test is used to determine how well someone thinks compared to other people their age. In general, IQ tests measure vocabulary knowledge, reasoning skills, the ability to use information, and how quickly someone thinks. IQ tests are often used in schools to determine if a child needs special services. Students whose IQ is much lower than students who are the same age might qualify for special education, whereas a student whose IQ is much higher might qualify for gifted services. IQ tests are also used to determine if someone is eligible for social security or disability benefits. An individual's score is compared to a larger sample of people the same age. For instance, a 17-year-old would be compared to other 17-year-olds. The average IQ score is 100, meaning that half of the population has an IQ of 100 or below and the other half is above 100. The majority of people in the United States fall within the average range which is from 85 to 115.

Q. Did you review any documents that suggested that Mr. Stevens' had been given an IQ test at some point when he was a child?

A. Yes, Mr. Stevens' IQ was tested when he was 12 years old and in the 5th grade. His verbal IQ score, which reflects his ability to use and understand words and language, was 83. His performance IQ, which is a measure of his visual and spatial skills, was very similar at 81. Visual and spatial skills allow us to look at something and understand it – for example, perceiving images in a mirror or being able to pack a suitcase or complete a puzzle. His overall IQ score, which is a measure of all the skills measured by the test, including the words that he knows, information that is commonly known in the community, and his ability to look at objects and perceive them accurately, was similar to his other two scores – it was 81.

What is important to understand is that all of these scores are below average –below what other people in the United States earns. His scores fall in the low average range, which is outside of the

range for intellectual disability. His score is too high to be classified as intellectually disabled. A score of 81 also falls in the borderline range, which is just above the cut off for intellectual disability. Regardless of which range you call it, only about 10% of the population scores this low on an IQ test.

Q. Doctor, tell us more about what an IQ score of 81 means.

A. Well, most people who take this test score somewhere between 85 and 115, with the majority of people scoring close to 100. When you look at all possible scores, 50% of people score below 100 and 50% above 100. He fell in the bottom 10th percentile meaning that 90% of the children his age scored higher than he did on this test when it was being made.

Q. Does a person's intelligence, as measured by an IQ test, affect the way a person functions at school? Meaning the person's ability to understand topics and participate in classes and interact with students and teachers?

A. Yes, a person's intelligence has a direct impact on a person's ability to learn and to function well in an educational setting. Mr. Stevens' score is in-line with his grades. His grades were low throughout all of his schooling with the majority of his grades being C's and some D's, but he also earned some F's in the more difficult courses. What is important to think about is that he was not tested or provided services until the 5th grade. Had he been tested earlier, I am certain that the teachers would have recognized that his intellectual skills were lower than most and that he could benefit from extra assistance.

Q. Is that sometimes referred to as special education?

A. Yes. In my opinion, because the teachers did not know the extent of his problems in learning they focused on trying to manage his behavior problems. It is not at all surprising that Mr. Stevens didn't behave well in the school setting. He was not doing well in his classes and it was clear that he couldn't keep up with his friends. Kids like this often give up on learning and spend their time disrupting the classroom. In his case, Mr. Stevens was getting in trouble in school for arguing with teachers, arguing with his peers, and using profanity in the classroom. Typically if kids do well and succeed in the classroom, their behavior improves. If they don't succeed, they tend to act out and get attention in a negative way. According to Mr. Stevens' mother and his 4th grade teacher, her son was socially promoted through the 7th grade and eventually quit school because his grades were so low.

Q. Did you give Mr. Stevens' an IQ test for the current evaluation?

A. I did. Mr. Stevens obtained a full scale IQ of 79 which is in the 8th percentile. That means that only 8% of individuals his age scored lower than he did at the time the test was made. In other words, 92% of his peers scored higher than he did on this test. Similar to prior testing, his full

scale IQ falls just above the range necessary for a diagnosis of intellectual disability in the borderline range.

Q. Tell us why intelligence is important to consider when you think about how he developed into the person that he is today, as well as the person that he was at the time of the crime.

A. Research has shown that doing very poorly in school, and doing so from an early age, is a risk factor for many things – such as substance abuse, adolescent criminal behavior, and school dropout. Basically, children who are not as "smart," if you will, have fewer opportunities in life, often have problems socializing with others, and often feel bad about themselves and these decrease the chances of success during teenage and adult years. With regard to opportunities, individuals like Mr. Stevens, do not typically succeed at sports or clubs such as chess or debate – participation in extracurricular activities such as sports and clubs are protective factors. Instead, people like Mr. Stevens will often stay to themselves, but some will join in with negative peer groups.

If you think about an IQ score as a measure of ability – which psychologists do – you have to consider that Mr. Stevens began school at a disadvantage. He was not as bright as other students and he had more difficulty solving problems and he didn't have the ability to weigh the pros and cons of his behavior in the same way that people with higher intelligence can.

APPENDIX G

Adaptive Behavior Manipulations

AB DEFICIT CONDITION

Q. Okay, let's move on to the second prong of the diagnosis of intellectual disability – adaptive behavior. Will you please tell the court what adaptive behavior refers to – I don't think that this term is something that most people know about.

A. Adaptive behavior is the ability of a person to function independently in the community and to do so on a day-to-day basis. Examples include the ability to dress and/or wash oneself, manage money, participate in conversations in a logical manner with an understanding of what is being said, making and maintaining friendships, and the like. The majority of people with intellectual disability will show some deficits in adaptive behavior and will require support or assistance in completing some daily activities. For example, an individual might need to be reminded to brush their teeth or need assistance using the stove. Others might have trouble following instructions that have multiple steps like checking the oil in the car or paying bills.

Q. Do all people with intellectual disability function the same way?

A. No. The more severe the level of intellectual disability the more deficits they will show and the more support they will need. The less severe the level of intellectual ability, the more independent they can be. In fact, many people with intellectual disability can function very well. In fact, so well that, unless you are very familiar with this person – such as a family member or friend – you wouldn't necessarily know that the person had intellectual disability.

Individuals with mild intellectual disability do not have the typical physical features associated with more severe cases: they do not speak with gestures and single word phrases and do not require constant supervision typical of more severe levels. Instead, for someone with a more mild level of disability you might expect to see some social immaturity and gullibility along with difficulties in planning and/or setting priorities. Often, individuals with mild intellectual disability will require support with tasks like managing finances, meal planning, and raising a family. Of note, individuals with low IQ scores are often able to function pretty well on their own in some areas, like self-care. Those skills, like brushing teeth and bathing, tend to be mastered early on in life. The deficits generally seen in individuals with IQs like Mr. Stevens include difficulties in budgeting, home maintenance, and meal planning. These individuals could not hold a job that required complex thinking or those that required a high level of independence. They can, however, do other jobs in custodial services, landscaping, construction or factory work, and other similar fields where the job is structured, easily understood, and carried out with supervision.

Q. Are deficits in adaptive behavior are required for a diagnosis of intellectual disability?

A. Yes they are.

Q. What do mental health professionals do to find out if a person has deficits in adaptive behavior?

A. Adaptive behavior is evaluated in a number of ways, primarily through record review, interviews with people who know the person very well, and in most cases, we have one or more people complete a rating form that tells us how well, in their experience, the person is able to carry out daily tasks. These are referred to as adaptive behavior instruments or tests. For this type of evaluation, you seek out people who knew the individual at different times during their life and in different contexts. Often teachers, family members, employers, and friends are good sources of information. These people would be interviewed and if they know the person well enough, they will be asked to complete a measure of adaptive behavior.

Q. Is there one test of adaptive behavior or does the evaluator have the option to choose from different tests?

A. There are many tests that can be used, but only a few of them are used to diagnose intellectual disability. The adaptive behavior measures are different from one another in some ways, but all of them require that someone with great knowledge of the person, rate that person on behaviors that are carried out by most people on a day-to-day basis. The questions are worded differently and the ratings are not always the same, but all of them assess multiple areas of functioning and these are generally the same for all measures.

Q. What do you mean when you say they assess multiple areas of functioning?

A. What I mean by that is that the instruments asks for ratings about their behavior in different settings and/or areas of functioning – for example within the work setting, school setting, or home setting – the main question that is asked is how well does the person complete a task without assistance from others. For example, some questions will be about the person's ability to take care of daily tasks related to personal hygiene, other questions will be pertinent to the person's ability to cook for themselves or take care of their home environment, still other skills relate to managing finances and banking. In addition to completing a measure of adaptive behavior, these people are also interviewed in order to obtain more detailed information.

Q. So these adaptive behavior skills, can they be categorized or grouped together in a meaningful way?

A. Yes. In fact, you can think of adaptive behavior skills as falling into three skill areas: conceptual, practical, and social skills.

Q. Please tell the court what you mean when you say conceptual skills.

A. Conceptual skills include communication, the ability to plan and strategize, and the application of academic skills. For example, the ability to apply simple math logic in calculating totals at the grocery store or determining how much money to give a cashier. Adults with conceptual skills deficits may have impaired short-term memory or poor priority setting skills.

Q. How about practical skills – what are those?

A. Practical skills include the ability to care for oneself and one's belongings. This would include the ability to respond appropriately to an injury or illness and adhere to a maintenance schedule on a car. Someone with deficits in the practical domain may need help to raise a family or prepare meals.

Q. And social skills?

A. Finally, social skills involve the ability to appropriately interact with others. Examples of social skills include demonstrating manners in one's interactions and effectively communicating one's needs or feelings about a situation. An individual with deficits in the social domain will likely be immature socially and may be at risk for manipulation by others.

Q. Doctor, as part of this evaluation, did you evaluate Mr. Stevens' adaptive functioning?

A. Yes I did.

Q. Please tell the court how you did that.

A. I relied primarily on records and information from Ms. Joanne Whitley, a woman who served as a mother-type figure for Mr. Stevens during his childhood. She completed a measure of adaptive behavior and I interviewed her for approximately three hours about her experiences with Mr. Stevens.

Q. What did you learn about Mr. Stevens?

A. Based on the information provided by Ms. Whitley, it became apparent that Mr. Stevens has demonstrated deficits in the conceptual domain since an early age. She reported that Mr. Stevens has always had trouble managing money and at the time of his arrest he struggled with knowing how to pay for things at a store. His biggest problem in this area was how much change he had to give to the cashier and because of this he tended to use paper money and receive change from the cashier. This was not a perfect answer to the problem because he also had difficulty knowing if he had been given the correct change.

Ms. Whitley also stated that Mr. Stevens consistently struggled with new situations when he had to follow a schedule. For example, if he was on a schedule for completing chores or taking care of maintenance on a car, he would often be unable to follow it. She also stated that he had trouble following instructions that were more than just a few steps in length. For example, Ms. Whitley would ask Mr. Stevens to help her with errands, such as going to the grocery store, but

he consistently forgot the grocery list at home or could not figure out where the items were in the grocery store or what brand to by, even if it was written out for him. She taught him how to check the oil on her car, but he was never able to check it on his own because he could not remember the steps to take. This report was consistent with school records which indicated he would function poorly in the classroom with tasks that were several steps long. Teachers noted that Mr. Stevens frequently required extra instruction in the classroom and repetition of directions to complete an assignment.

Mr. Stevens also struggled quite a bit at home. He required assistance to do things on his own. Ms. Whitley gave the example that he needed to be reminded to pick up after himself and consistently do his laundry. Ms. Whitley indicated that he was eventually able to do these things on his own eventually, but he did require extra reminders. Based on her reports, it seemed as though Mr. Stevens later learned to do these things independently, but that he did need extra help until about age 15.

Mr. Stevens had difficulty when he was enrolled in school as well. His teachers indicated he required near constant supervision to stay on task and keep up with the other kids because he just had so much difficulty following instructions. Records consistently described Mr. Stevens as off task and behind the other kids in terms of following instructions. His lack of academic progress in school was also well documented, as I have mentioned before – he failed the 5th grade. His 5th grade teacher filled out a list of concerns about Mr. Stevens and recommended an evaluation due to her concerns about his lack of progress and need for additional help.

Ms. Whitley also reported that Mr. Stevens had trouble communicating with others and that this was true in multiple situations. She noted that it was not uncommon for him to misunderstand others and he also had trouble communicating his needs and/or thoughts with people. Her reports suggested that Mr. Stevens would frequently acquiesce – or simply say “yes” or agree in order to give the façade that he understood. It seems as though Mr. Stevens was able to mask of his difficulties in this manner; though it became apparent when he was unable to complete the tasks required of him that he did not understand.

Taken together, the difficulties described by Ms. Whitley appeared to significantly impact Mr. Stevens in social, academic and vocational settings. Ms. Whitley’s reports on the adaptive behavior measure were also consistent with deficits in his ability to apply academic skills to real life situations, which is consistent with deficits in the conceptual domain.

Q. Doctor, did you do any testing that extended your understanding of Mr. Stevens’ abilities within the academic realm?

A. Yes I did. I administered an achievement test to evaluate how Mr. Stevens was functioning, academically, as compared to other people his age.

Q. How do people interpret the scores on these tests?

A. There are a number of ways, but most people will look at either grade equivalent scores – how the person compares to others in the same grade or age equivalent scores which is comparing that persons scores with people of the same age.

Q. How did you interpret the scores?

A. I looked at grade equivalent scores which allow you to see if the person is functioning similar to his peers in the same grade, above his peers, or below his peers. This test that I gave him provides scores for four areas that are usually assessed in the school environment – these are math, reading, writing, and language. Given that the test is academic in nature, the data obtained from this test allows for a more in depth look at his abilities as related to the conceptual domain.

Q. How did Mr. Stevens do on this test?

A. He did very poorly. By my assessment, Mr. Stevens reads at about a third grade level. His writing and language skills are at the fourth grade level and his spelling is at the beginning of the fifth grade level. His math skills fell at about the third grade level. What is important here is that his problems in functioning as reported by Ms. Whitley are consistent with these test scores. In order to apply skills, it is important that you have mastered the skills in an academic setting. As seen by these scores, his mastery of academic skills is far below what we would expect. In other words, it makes sense that he has these difficulties as his level of academic achievement is low.

Q. It would be rare to find someone with a low IQ without academic deficits – wouldn't it?

A. Well, people with mild intellectual disability generally develop reading skills that fall between the 3rd and 6th grade levels. A 6th grade reading level is considered to be functionally literate, but this level is quite low and makes living on a day-to-day basis much harder than it is for others. For instance, the ability to read an informational or instructional packet typically requires literacy that is above the 6th grade, but it also requires the ability to understand the meaning of the information and to apply it to their lives. This is what is important when you are looking at deficits in abilities that are related to academic skills. Someone may be able to complete some math problems that are at the 6th grade, but struggle to understand how to provide the correct amount of money to a cashier or check how much change they should receive in return. Once again, the issue is whether his deficits create an impediment to functioning in the community on a day-to-day basis. In Mr. Stevens' case, it appears that his ability to apply academic knowledge in his daily life is impaired.

Q. What conclusions did you reach regarding adaptive behavior in the case of Mr. Stevens?

A. Based on a review of records, achievement testing, and information provided to me by Ms. Whitley, I came to the conclusion that, in my opinion, Mr. Stevens demonstrates significant deficits in the conceptual domain of functioning. Mr. Stevens does not have the ability to function in the way that most individuals do. He has required a great deal of assistance to

function on his own and this was well documented in school records and further reported by other sources including Ms. Whitley. Mr. Stevens' difficulty caring for himself and his belongings is consistent with deficits in the conceptual domain. His lack of academic progress and difficulty applying academic skills to real life, for example – in making correct change for a cashier – are other examples of areas where he has struggled.

Q. Did Mr. Stevens show deficits in either the social or practical areas of functioning?

A. Yes, he did show some deficits in functioning in the social realm, but they didn't impact his functioning as severely as his deficits in the conceptual domain. For example, although he struggles to communicate with others and has been described as immature, it does not seem as though his difficulties here have negative impacted him in such a way that he cannot function on his own. In other words, he was able to compensate for these difficulties and function like his peers.

Q. What about the practical domain?

A. No. He doesn't demonstrate deficits in the practical domain. From what I learned from the people that I spoke with about the case, he learned practical skills more slowly than did others and didn't master them to the degree as others did, but his problems were never so bad that he couldn't function.

Q. Doctor, in order to qualify for a diagnosis of intellectual disability, does a person have to show significant deficits in all three areas?

A. No.

Q. How many satisfy the criteria?

A. One.

Q. And in this case, the case of Mr. Stevens, how many satisfied this requirement?

A. Two.

Q. So the results of your evaluation led you to the opinion that he does meet the adaptive behavior criteria required for the diagnosis of intellectual disability?

A. Yes.

Q. Okay, moving on to…

NO AB DEFICIT CONDITION

Q. Okay, let's move on to the second part of the diagnosis of intellectual disability – adaptive behavior. Will you please tell the court what adaptive behavior refers to – I don't think that this term is something that most people know about.

A. Adaptive behavior is the ability of a person to function independently in the community and to do so on a day-to-day basis. Examples include the ability to dress and/or wash oneself, manage money, participate in conversations in a logical manner with an understanding of what is being said, making and maintaining friendships, and the like. The majority of people with intellectual disability will show some deficits in adaptive behavior and will require support or assistance in completing some daily activities. For example, an individual might need to be reminded to brush their teeth or need assistance using the stove. Others might have trouble following instructions that have multiple steps like checking the oil in the car or paying bills.

Q. Do all people with intellectual disability function the same way?

A. No. The more severe the level of intellectual disability the more deficits they will show and the more support they will need. The less severe the level of intellectual ability, the more independent they can be. In fact, many people with intellectual disability can function very well. In fact, so well that, unless you are very familiar with this person – such as a family member or friend – you wouldn't necessarily know that the person had intellectual disability.

Individuals with mild intellectual disability do not have the typical physical features associated with more severe cases: they do not speak with gestures and single word phrases and do not require constant supervision typical of more severe levels. Instead, for someone with a more mild level of disability you might expect to see some social immaturity and gullibility along with difficulties in planning and/or setting priorities. Often, individuals with mild intellectual disability will require support with tasks like managing finances, meal planning, and raising a family. Of note, individuals with low IQ scores are often able to function pretty well on their own in some areas, like self-care. Those skills, like brushing teeth and bathing, tend to be mastered early on in life. The deficits generally seen in individuals with IQs like Mr. Stevens include difficulties in budgeting, home maintenance, and meal planning. These individuals could not hold a job that required complex thinking or those that required a high level of independence. They can, however, do other jobs in custodial services, landscaping, construction or factory work, and other similar fields where the job is structured, easily understood, and carried out with supervision.

Q. Are deficits in adaptive behavior are required for a diagnosis of intellectual disability?

A. Yes they are.

Q. What do mental health professionals do to find out if a person has deficits in adaptive behavior?

A. Adaptive behavior is evaluated in a number of ways, primarily through record review, interviews with people who know the person very well, and in most cases, we have one or more people complete a rating form that tells us how well, in their experience, the person is able to carry out daily tasks. These are referred to as adaptive behavior instruments or tests. For this type of evaluation, you seek out people who knew the individual at different times during their life and in different contexts. Often teachers, family members, employers, and friends are good sources of information. These people would be interviewed and if they know the person well enough, they will be asked to complete a measure of adaptive behavior.

Q. Is there one test of adaptive behavior or does the evaluator have the option to choose from different tests?

A. There are many tests that can be used, but only a few of them are used to diagnose intellectual disability. The adaptive behavior measures are different from one another in some ways, but all of them require that someone with great knowledge of the person, rate that person on behaviors that are carried out by most people on a day-to-day basis. The questions are worded differently and the ratings are not always the same, but all of them assess multiple areas of functioning and these are generally the same for all measures.

Q. What do you mean when you say they assess multiple areas of functioning?

A. What I mean by that is that the instruments asks for ratings about their behavior in different settings and/or areas of functioning – for example within the work setting, school setting, or home setting – the main question that is asked is how well does the person complete a task without assistance from others. For example, some questions will be about the person's ability to take care of daily tasks related to personal hygiene, other questions will be pertinent to the person's ability to cook for themselves or take care of their home environment, still other skills relate to managing finances and banking. In addition to completing a measure of adaptive behavior, these people are also interviewed in order to obtain more detailed information.

Q. So these adaptive behavior skills, can they be categorized or grouped together in a meaningful way?

A. Yes. In fact, you can think of adaptive behavior skills as falling into three skill areas: conceptual, practical, and social skills.

Q. Please tell the court what you mean when you say conceptual skills.

A. Conceptual skills include communication, the ability to plan and strategize, and the application of academic skills. For example, the ability to apply simple math logic in calculating

totals at the grocery store or determining how much money to give a cashier. Adults with conceptual skills deficits may have impaired short-term memory or poor priority setting skills.

Q. How about practical skills – what are those?

A. Practical skills include the ability to care for oneself and one's belongings. This would include the ability to respond appropriately to an injury or illness and adhere to a maintenance schedule on a car. Someone with deficits in the practical domain may need help to raise a family or prepare meals.

Q. And social skills?

A. Finally, social skills involve the ability to appropriately interact with others. Examples of social skills include demonstrating manners in one's interactions and effectively communicating one's needs or feelings about a situation. An individual with deficits in the social domain will likely be immature socially and may be at risk for manipulation by others.

Q. Doctor, as part of this evaluation, did you evaluate Mr. Stevens' adaptive functioning?

A. Yes I did.

Q. Please tell the court how you did that.

A. I relied primarily on records and information from Ms. Joanne Whitley, a woman who served as a mother-type figure for Mr. Stevens during his childhood. She completed a measure of adaptive behavior and I interviewed her for approximately three hours about her experiences with Mr. Stevens.

Q. What did you learn about Mr. Stevens?

A. Based on the information provided by Ms. Whitley, it became apparent that although Mr. Stevens had some difficulty functioning independently, he did not demonstrate severe enough deficits to qualify for deficits in adaptive behavior. She reported that Mr. Stevens had trouble managing money and knowing how to pay for things at a store when he was younger. His biggest problem in this area was how much change he had to give to the cashier and because of this he tended to use paper money and receive change from the cashier. This was not a perfect answer to the problem because he also had difficulty knowing if he had been given the correct change. Mr. Stevens struggled with this quite a bit during childhood but was eventually able to learn how to do so on his own and without assistance.

Although he did require some additional instruction in learning tasks, Mr. Stevens was able to catch up over time and with repetition. Ms. Whitley reported that he would initially struggle with new situations when he had to follow a schedule. He had to be reminded and taught how to complete chores and follow car maintenance schedules before he was able to perform these tasks on his own. Ms. Whitley spent extra time teaching him how to help her with errands, such as

going to the grocery store. Initially he would forget the list or be unable to locate items, but over time he was able to buy groceries on his own. She taught him how to check the oil on her car, and after watching her do it and doing it under her supervision, he was able to check the oil on his own. This report was consistent with school records that indicated he would initially function poorly in the classroom with tasks that were more than just a couple steps. With help and practice, however, he was typically able to perform the necessary tasks independently.

Mr. Stevens struggled quite a bit at home. He required assistance to do things on his own. Ms. Whitley gave the example that he needed to be reminded to pick up after himself and to consistently do his laundry. Ms. Whitley indicated that he was eventually able to do these things on his own, but he did require extra reminders. Based on her reports, it seemed as though Mr. Stevens later learned to do these things independently, but that he did need extra help until about age 15.

As mentioned previously, Mr. Stevens had difficulty when he was enrolled in school. His teachers indicated he required regular supervision to stay on task and keep up with the other kids because of his difficulty following instructions. Records consistently described Mr. Stevens as off task and behind the other kids in terms of following instructions. His lack of academic progress in school was also well documented, as I have mentioned before, he failed the 5th grade. His 5th grade teacher filled out a list of concerns about Mr. Stevens and recommended an evaluation due to her concerns about his lack of progress and need for additional help.

Ms. Whitley denied that Mr. Stevens had trouble with communication with others. He did not appear to misunderstand others nor did he have trouble communicating his thoughts or feelings with other people. Despite this, Mr. Stevens did seem to acquiesce – or simply say “yes” or agree in order to give the façade that he understand when he really did not. It seems as though he was able to mask some of his difficulties related to his intelligence in this manner.

Taken together, it does not appear that his adaptive behavior deficits significantly impacted his functioning in social, academic, or vocational settings. That is not to say that he did not struggle and have difficulty, but it seems with enough instruction and assistance at the outset, Mr. Stevens was able to perform adequately on his own. Therefore, based on the available information, it appears his difficulties did not rise to the level that he demonstrated true adaptive behavior deficits we would expect to see in an individual with intellectual disability.

Q. Doctor, did you do any testing that extended your understanding of Mr. Stevens’ abilities within the academic realm?

A. Yes I did. I administered an achievement test to evaluate how Mr. Stevens was functioning, academically, as compared to other people his age.

Q. How do people interpret the scores on these tests?

A. There are a number of ways, but most people will look at either grade equivalent scores – how the person compares to others in the same grade or age equivalent scores which is comparing that persons scores with people of the same age.

Q. How did you interpret the scores?

A. I looked at grade equivalent scores which allow you to see if the person is functioning similar to his peers in the same grade, above his peers, or below his peers. This test that I gave him provides scores for four areas that are usually assessed in the school environment – these are math, reading, writing, and language. Given that the test is academic in nature, the data obtained from this test allows for a more in depth look at his abilities as related to the conceptual domain.

Q. How did Mr. Stevens do on this test?

R. He did very poorly. By my assessment, Mr. Stevens reads at about a third grade level. His writing and language skills are at the fourth grade level and his spelling is at the beginning of the fifth grade level. His math skills fell at about the third grade level. What is important here is that his problems in functioning, as mentioned previously, are consistent with these test scores. In order to apply skills, it is important that you have mastered the skills in an academic setting. As seen by these scores, his mastery of academic skills is far below what we would expect. In other words, it makes sense that he has these difficulties as his level of academic achievement is low.

Q. It would be rare to find someone with a low IQ without academic deficits – wouldn't it?

A. Well, people with mild intellectual disability generally develop reading skills that fall between the 3rd and 6th grade levels. A 6th grade reading level is considered to be functionally literate, but this level is quite low and makes living on a day-to-day basis much harder than it is for others. For instance, the ability to read an informational or instructional packet typically requires literacy that is above the 6th grade, but it also requires the ability to understand the meaning of the information and to apply it to their lives. This is what is important when you are looking at deficits in abilities that are related to academic skills. Someone may be able to complete some math problems that are at the 6th grade, but struggle to understand how to provide the correct amount of money to a cashier or check how much change they should receive in return. Once again, the issue is whether his deficits create an impediment to functioning in the community on a day-to-day basis. In Mr. Stevens' case, it appears that his ability to apply academic knowledge in his daily life is not impaired to the degree that would qualify for deficits in adaptive behavior.

Q. What conclusions did you reach regarding adaptive behavior in the case of Mr. Stevens?

A. Based on a review of records, achievement testing, and information provided to me by Ms. Whitley, I came to the conclusion that, in my opinion, Mr. Stevens does not demonstrate significant deficits in adaptive behavior. Although Mr. Stevens required extra assistance and was

slower to catch on than others, he does appear to have the ability to function the way that most other individuals do. He required some initial assistance to function on his own and this was well documented in school records and further reported by other sources including Ms. Whitley. His lack of academic progress and difficulty applying academic skills to real life, for example – in making correct change for a cashier – are examples of areas where he struggled and required additional instruction. That said, his difficulties did not rise to the level of impairment that would qualify for deficits in adaptive behavior.

Q. Did Mr. Stevens show deficits in either the social or practical areas of functioning?

A. Yes, he did show some deficits in functioning in the social realm, but they didn't impact his functioning so severely that he would meet criteria for deficits in these areas. For example, although he struggles to communicate with others and has been described as immature, it does not seem as though his difficulties here have negative impacted him in such a way that he cannot function on his own. In other words, he was able to compensate for these difficulties and function like his peers.

Q. What about the practical domain?

A. No. He doesn't demonstrate deficits in the practical domain. From what I learned from the people that I spoke with about the case, he learned practical skills more slowly than did others and didn't master them to the degree as others did, but his problems were never so bad that he couldn't function.

Q. Doctor, in order to qualify for a diagnosis of intellectual disability, does a person have to show significant deficits in all three areas?

A. No.

Q. How many satisfy the criteria?

A. One.

Q. And in this case, the case of Mr. Stevens, how many satisfied this requirement?

A. None.

Q. So the results of your evaluation led you to the opinion that Mr. Stevens does not meet the adaptive behavior criteria required for the diagnosis of intellectual disability?

A. No. He does not meet the adaptive behavior criteria.

Q. Okay, moving on to…

APPENDIX H

Description of the Crime

Mr. Michael Stevens went out "looking for some robberies" in Pickens County, Alabama in May 2012. During the early evening hours, Mr. Stevens decided on a victim – a young woman who was stranded on the side of the road after her car broke down. Mr. Stevens stopped when he saw her alone by her broken down car on the side of a county road. According to testimony given during the guilt-phase of the trial, Mr. Stevens approached the woman and shot her twice in the chest and once in the head. Afterward, he emptied her pockets and stole the purse out of the car. Mr. Stevens was tried and found to be guilty of murder in the first degree for which he could be sentenced to death.

Witnesses testified that they saw Mr. Stevens flee from the scene of the crime. Mr. Stevens went back to his friend's neighborhood where he changed his blood stained clothes and returned the gun to the closet where it was kept. According to friends of Mr. Stevens, he was joking around and bragging about the money he received from his robbery. According to one witness, Mr. Stevens tried to establish an alibi for the night of the crime and asked one of his friends to tell the police that they had been together that evening. Some of Mr. Stevens' friends testified that he was excited that he pulled off the crime because he thought that he would be put to a higher rank in the gang he was involved in.

APPENDIX I

Pre-Deliberation Research Packet

Participant Number: ____

Jury Number: ____

Based on the testimony you have just heard, which of the following sentences would you vote for in the case of *Michael Stevens v. State*?

Circle one:

Life Imprisonment without the Possibility of Parole **Death Penalty**

Please rate the confidence you have in your decision. In other words, how confident are you with the decision to sentence Mr. Stevens to either life in prison or death?

1 2 3 4 5 6 7 8 9 10

(not at all confident) (very confident)

According to the expert who testified, did Michael Stevens have impaired intellectual functioning (i.e., IQ) consistent with intellectual disability (ID)? *Please circle.*

Yes No

Did Michael Stevens demonstrate deficits in adaptive behavior severe enough to qualify for a diagnosis of intellectual disability (ID)? *Please circle.*

Yes No

Please rate each of the following factors in terms of how mitigating or aggravating you view it to be. Recall that mitigating factors are those that support a sentence of life imprisonment whereas aggravating factors are those that support a sentence of death.

Low intellectual functioning (IQ).

1 2 3 4 5 6 7 8 9 10

(Completely mitigating) (Completely aggravating)

Adaptive behavior. *(Trouble managing money, difficulty with schedules and multistep instructions, communication difficulty, poor academic skills).*

1 2 3 4 5 6 7 8 9 10

(Completely mitigating) (Completely aggravating)

Childhood onset psychological disorder. *(Diagnosed with ADHD at age 7, hyperactive and impulsive, uncooperative with treatment).*

1 2 3 4 5 6 7 8 9 10

(Completely mitigating) (Completely aggravating)

Exposure to crime and violence. *(Exposure to violence and crime in other neighborhoods, gang activity, modeling of criminal behavior and violence).*

1 2 3 4 5 6 7 8 9 10
(Completely mitigating) (Completely aggravating)

Illicit substance use. *(Drug and alcohol use modeled by adults around him, drug and alcohol use since an early age).*

1 2 3 4 5 6 7 8 9 10
(Completely mitigating) (Completely aggravating)

Parenting. *(Mother was absent because she was working as a nurse to provide for family, unsupervised).*

1 2 3 4 5 6 7 8 9 10
(Completely mitigating) (Completely aggravating)

Father absence. *(Minimal contact and involvement of father).*

1 2 3 4 5 6 7 8 9 10
(Completely mitigating) (Completely aggravating)

Failure to provide supervision and guidance. *(Mother was often gone and kids were unsupervised, absence of positive role models).*

1 2 3 4 5 6 7 8 9 10
(Completely mitigating) (Completely aggravating)

Youthfulness. *(Age 19 at time of offense, brain not fully developed).*

1 2 3 4 5 6 7 8 9 10
(Completely mitigating) (Completely aggravating)

The murder was committed during a robbery, the aggravating factor agreed upon during the guilt-phase court.

1 2 3 4 5 6 7 8 9 10
(Completely mitigating) (Completely aggravating)

Please circle all of the following factors that you ***<u>considered to be mitigating</u>*** *in this case. In other words, which of the following factors led you to think that Mr. Stevens might be more deserving of a sentence of LWOP than death. Of note, it is ok to view some factors to be mitigating, yet still decide that death is more appropriate because of how you weighed mitigating versus aggravating factors. Similarly, it you only found that there were only a few factors that you considered to be mitigating, but you found them to be quite powerful, you can vote for LWOP over death.*

- Low intellectual functioning (IQ)
- Adaptive behavior.
- Childhood onset psychological disorder (ADHD).
- Exposure to crime and violence.
- Illicit substance use.
- Parenting.
- Father absence.

- Failure to provide supervision and guidance.
- Youthfulness.

In your opinion, is Mr. Stevens intellectually disabled? Yes No

Why or why not?

_____

_____

_____

_____

_____

APPENDIX J

Post-Deliberation Research Packet

Jury Number: _____
Participant Number: ____

Based on the testimony you heard and the deliberation you just participated in, which of the following sentences would you choose for Mr. Michael Stevens?

Circle one:

Life Imprisonment without the Possibility of Parole **Death Penalty**

Please rate the confidence you have in your decision. In other words, how confident are you with the decision to sentence Mr. Stevens to either life in prison or death?

1 2 3 4 5 6
(not at all confident) (very confident)

Did your sentence change from pre-deliberation ratings? Yes No

Why or why not? _____
_____
_____

Please rate each of the following factors in terms of how mitigating or aggravating you view it to be. Recall that mitigating factors are those that support a sentence of life imprisonment whereas aggravating factors are those that support a sentence of death.

Low intellectual functioning (IQ).

1 2 3 4 5 6 7 8 9 10
(Completely mitigating) (Completely aggravating)

Adaptive behavior. *(Trouble managing money, difficulty with schedules and multistep instructions, communication difficulty, poor academic skills).*

1 2 3 4 5 6 7 8 9 10
(Completely mitigating) (Completely aggravating)

Childhood onset psychological disorder. *(Diagnosed with ADHD at age 7, hyperactive and impulsive, uncooperative with treatment).*

1 2 3 4 5 6 7 8 9 10
(Completely mitigating) (Completely aggravating)

Exposure to crime and violence. *(Exposure to violence and crime in other neighborhoods, gang activity, modeling of criminal behavior and violence).*

1 2 3 4 5 6 7 8 9 10

(Completely mitigating) (Completely aggravating)

Illicit substance use. *(Drug and alcohol use modeled by adults around him, drug and alcohol use since an early age).*

1 2 3 4 5 6 7 8 9 10

(Completely mitigating) (Completely aggravating)

Parenting. *(Mother was absent because she was working as a nurse to provide for family, unsupervised).*

1 2 3 4 5 6 7 8 9 10

(Completely mitigating) (Completely aggravating)

Father absence. *(Minimal contact and involvement of father).*

1 2 3 4 5 6 7 8 9 10

(Completely mitigating) (Completely aggravating)

Failure to provide supervision and guidance. *(Mother was often gone and kids were unsupervised, absence of positive role models).*

1 2 3 4 5 6 7 8 9 10

(Completely mitigating) (Completely aggravating)

Youthfulness. *(Age 19 at time of offense, brain not fully developed).*

1 2 3 4 5 6 7 8 9 10

(Completely mitigating) (Completely aggravating)

The murder was committed during a robbery, the aggravating factor agreed upon during the guilt-phase court.

1 2 3 4 5 6 7 8 9 10

(Completely mitigating) (Completely aggravating)

Instructions: *Before you move on, write "I read the instructions" at the top of this page.*

Please circle all of the following factors that you ***<u>considered to be mitigating</u>*** *in this case. In other words, which of the following factors led you to think that Mr. Stevens might be more deserving of a sentence of LWOP than death. Of note, it is ok to view some factors to be mitigating, yet still decide that death is more appropriate because of how you weighed mitigating versus aggravating factors. Similarly, it you only found that there were only a few factors that you considered to be mitigating, but you found them to be quite powerful, you can vote for LWOP over death.*

- Low intellectual functioning (IQ)
- Adaptive behavior.
- Childhood onset psychological disorder (ADHD).
- Exposure to crime and violence.
- Illicit substance use.
- Parenting.

- Father absence.
- Failure to provide supervision and guidance.
- Youthfulness.

In your opinion, is Mr. Stevens intellectually disabled? Yes No
Why or why not?

_____
_____
_____
_____
_____

APPENDIX K

Revised Legal Authoritarianism Scale (RLAQ)

Please rate your agreement with each of the following statements. Do not skip any items.

1. Unfair treatment of underprivileged groups and classes is the chief cause of crime.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

2. Too many obviously guilty persons escape punishment because of legal technicalities.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

3. Evidence illegally obtained should be admissible in court if such evidence is the only way of obtaining a conviction.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

4. Search warrants should clearly specify the person or things to be seized.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

5. No one should be convicted of a crime on the basis of circumstantial evidence, no matter how strong such evidence is.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

6. There is no need in a criminal case for the accused to prove his innocence beyond a reasonable doubt.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

7. Any person who resists arrest commits a crime.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

8. When determining a person's guilt or innocence, the existence of a prior arrest record should be considered.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

9. Wiretapping by anyone for any reason should be completely illegal.

1	2	3	4	5	6

Strongly Disagree	Strongly Agree

10. Defendants in a criminal case should be required to take the witness stand.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

11. All too often, minority group members do not get fair trials.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

12. Because of the oppression and persecution minority group members suffer, they deserve leniency and special treatment in the courts.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

13. Citizens need to be protected against excess police power as well as against criminals.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

14. It is better for society that several guilty men be freed than one innocent one wrongfully imprisoned.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

15. Accused persons should be required to take lie-detector tests.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

16. When there is a “hung” jury in a criminal case, the defendant should always be freed and the indictment dismissed.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

17. A society with true freedom and equality for *all* would have very little crime.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

18. It is moral and ethical for a lawyer to represent a defendant in a criminal case even when he believes his client is guilty.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

19. Police should be allowed to arrest and question suspicious looking persons to determine whether they have been up to something illegal.

1	2	3	4	5	6

Strongly Disagree	Strongly Agree

20. The law coddles criminals to the detriment of society.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

21. The freedom of society is endangered as much by overzealous law enforcement as by the acts of individual criminals.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

22. In the long run, liberty is more important than order.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

23. Upstanding citizens have nothing to fear from the police.

1	2	3	4	5	6
Strongly Disagree					Strongly Agree

APPENDIX L

Death Penalty Attitudes Questionnaire

Please rate your agreement with each of the following statements. Do not skip any items.

1. I think the death penalty is necessary.
   1 2 3 4 5 6 7 8 9
   (strongly disagree) (strongly agree)

2. It is immoral for society to take a life regardless of the crime the individual has committed.
   1 2 3 4 5 6 7 8 9
   (strongly disagree) (strongly agree)

3. No matter what crime a person has committed executing them is a cruel punishment.
   1 2 3 4 5 6 7 8 9
   (strongly disagree) (strongly agree)

4. The death penalty should be used more often than it is.
   1 2 3 4 5 6 7 8 9
   (strongly disagree) (strongly agree)

5. The desire for revenge is a legitimate reason for favoring the death penalty.
   1 2 3 4 5 6 7 8 9
   (strongly disagree) (strongly agree)

6. Society has a right to get revenge when murder has been committed.
   1 2 3 4 5 6 7 8 9
   (strongly disagree) (strongly agree)

7. There are some murderers whose death would give me a sense of personal satisfaction.
   1 2 3 4 5 6 7 8 9
   (strongly disagree) (strongly agree)

8. The death penalty is the just way to compensate the victim's family for some murders.
   1 2 3 4 5 6 7 8 9
   (strongly disagree) (strongly agree)

9. The death penalty does not deter murderers.
   1 2 3 4 5 6 7 8 9
   (strongly disagree) (strongly agree)

10. The death penalty makes criminals think twice before committing murder.
    1 2 3 4 5 6 7 8 9
    (strongly disagree) (strongly agree)

11. Executing a person for premeditated murder discourages others from committing that crime in the future.

1 2 3 4 5 6 7 8 9
(strongly disagree) (strongly agree)

12. It is more cost efficient to sentence a murderer to death rather than life imprisonment.
1 2 3 4 5 6 7 8 9
(strongly disagree) (strongly agree)

13. Executing a murderer is less expensive than keeping him in jail for the rest of his life.
1 2 3 4 5 6 7 8 9
(strongly disagree) (strongly agree)

14. Even when a murdered gets a sentence of life without parole, he usually gets out on parole.
1 2 3 4 5 6 7 8 9
(strongly disagree) (strongly agree)

15. There is no such thing as a sentence that truly means “life without parole.”
1 2 3 4 5 6 7 8 9
(strongly disagree) (strongly agree)

APPENDIX M

Statement of Seriousness of the Research

In this study, you will be asked to play the role of a juror. It is important that you take this role seriously as the results of this study could be used to inform court cases in the future. The materials in this project have been adapted from an actual capital case. The results of this data will provide us with information about how actual capital decisions are made. It is possible that the results of this study could lead us to improve this type of decision-making in the future. It is important that you take your role as a juror seriously.

APPENDIX N

Deliberation Instructions

Dismissal of participants 7 and above to alternative location with second researcher:
Participants one through six will remain in this location for the next portion of the study. Participants assigned a number greater than seven will be moved to another location to complete the study. At this time, participants seven and above can exit with _____ (researcher name). Participants one through six, please arrange your seats so you are in numerical order in a circle with your participant number placed in front of you.

Instructions read to participants 1-6, the deliberating group:
We are now entering the deliberation phase of the study where you will discuss your decision-making and arrive at a unanimous decision with your peers. You will spend between 30 minutes and one hour deliberating, or deciding as a group, the sentence for the defendant, Mr. Stevens. All of you must agree on the sentence you chose for Mr. Stevens. Deliberations can end once you have reached a unanimous decision AND 30 minutes have elapsed. Participant number one will serve as the foreperson for this phase of the study.

Although a researcher will be in the room during your discussions, you are not permitted to ask them questions. Your discussions will be audio-recorded. Please refer to the informed consent document regarding the use of these recordings; as a reminder, your discussions will not be traced back to your identifying information.

Instructions read to participants 7 and above:
We are now entering the deliberation phase of the study where you will discuss your decision-making and arrive at a unanimous decision with your peers. You will spend between 30 minutes and one hour deliberating, or deciding as a group, the sentence for the defendant, Mr. Stevens. You all must agree on the sentence you chose for Mr. Stevens. Deliberations can end once you have reached a unanimous decision AND 30 minutes have elapsed. Participant number seven will serve as the foreperson for this phase of the study.

Although a researcher will be in the room during your discussions, you are not permitted to ask them questions.

In the case in which only one participant composes the second group:
Please independently think about the sentence you decided on for approximately ten minutes. Consider which factors influenced your decision and how confident you are in your decision. At the end of ten minutes, the researcher will provide you with additional materials to complete the study.

APPENDIX O

Qualitative Deliberation Form

Jury Number: ____

Length of Deliberation (in minutes): _____

Number of female participants: ____

Number of Caucasian participants: _____

Characteristics of Foreperson: Male Female
Caucasian Minority
Life Death
Other (including age, education, etc.):

Initial Vote: _____ (life) to _____ (death)

First Explicit Shift **to** (check one): _____ Life _____ Death

Deliberation Style (check one): _____ Verdict-Driven _____ Evidence-Driven

Final Sentence (check one): _____ Life _____ Death

Notes on Contributions: _____
_____
_____
_____
_____
_____

Record order of contributions:
_____
_____
_____
_____
_____
_____

APPENDIX P

Standard Deliberation Prompts

If the mock jury does not deliberate, the researcher will say the following: **Please discuss how you arrived at your sentence for Mr. Stevens with your peers. You may discuss any aspect you thought was important in arriving at your decision.**

In the case of an initial unanimous verdict, the researcher will say the following: **Although you have already reading a unanimous verdict, please discuss with each other how you arrived at this conclusion. You may want to discuss which aspects of the testimony influenced your decision or which factors were irrelevant to you.**

If the mock jury halts deliberations, the researcher will say the following: **Please continue to discuss your decision-making process for an additional ___ minutes until the time limit has elapsed.**

If a member of the mock jury asks a question about their task, the researcher will say the following: **Please refer to the written instructions in front of you. I am unable to provide any additional information about your task.**

APPENDIX Q
IRB Approval

October 24, 2014

Office for Research

Institutional Review Board for the Protection of Human Subjects

THE UNIVERSITY OF ALABAMA RESEARCH

Mary Wood
Department of Psychology
College of Arts & Sciences
Box 870348

Re: IRB # 14-OR-360, "Beyond Atkins: How do the Prongs Perform During Sentencing?"

Dear Ms. Wood:

The University of Alabama Institutional Review Board has granted approval for your proposed research.

Your application has been given expedited approval according to 45 CFR part 46. You have also been granted the requested waiver of informed consent. Approval has been given under expedited review category 7 as outlined below:

(7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

Your application will expire on October 22, 2015. If your research will continue beyond this date, please complete the relevant portions of the IRB Renewal Application. If you wish to modify the application, please complete the Modification of an Approved Protocol form. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants. When the study closes, please complete the Request for Study Closure form.

Should you need to submit any further correspondence regarding this proposal, please include the above application number.

Good luck with your research.

Sincerely,

Stuart Usdan, Ph.D.
Chair, Non-Medical IRB
The University of Alabama

358 Rose Administration Building
Box 870127
Tuscaloosa, Alabama 35487-0127
(205) 348-8461
FAX (205) 348-7189
TOLL FREE (877) 820-3066

IRB Project #:

UNIVERSITY OF ALABAMA
INSTITUTIONAL REVIEW BOARD FOR THE PROTECTION OF HUMAN SUBJECTS
REQUEST FOR APPROVAL OF RESEARCH INVOLVING HUMAN SUBJECTS

I. Identifying information

	Principal Investigator	Second Investigator	Third Investigator
Names:	Mary E. Wood	Karen L. Salekin	
Department:	Psychology	Psychology	
College:	Arts and Sciences	Arts and Sciences	
University:	UA	UA	
Address:	Box 870348	Box 870348	
Telephone:	205-277-8005		
FAX:			
E-mail:	mschnorff@crimson.ua.edu	ksalekin.ua.edu	

Title of Research Project: Beyond Atkins: How do the Prongs Perform During Sentencing?

Date Submitted: 09/09/2014
Funding Source: N/A

Type of Proposal	☒ New	☐ Revision	☐ Renewal Please attach a renewal application	☐ Completed	☐ Exempt
			Please attach a continuing review of studies form		
		Please enter the original IRB # at the top of the page			

UA faculty or staff member signature: _____

II. NOTIFICATION OF IRB ACTION (to be completed by IRB):
Type of Review: _____ Full board X Expedited

IRB Action:
___ Rejected Date: _____
___ Tabled Pending Revisions Date: _____
___ Approved Pending Revisions Date: _____
X Approved-this proposal complies with University and federal regulations for the protection of human subjects.

Approval is effective until the following date: 10-22-15
Items approved: X Research protocol (dated 10-23-14)
X Informed consent (dated 10-23-14)
___ Recruitment materials (dated _____)
___ Other (dated _____)

Approval signature _____ Date _____

PARTICIPANT INFORMATION SHEET

Study title: Beyond *Atkins*: How do the Prongs Perform during Sentencing?

Principal Investigator: Mary E. Wood, M.A., Graduate Student

Faculty Supervisor: Karen L. Salekin, Ph.D.

What is this study about? You are being asked to take part in a research study. This study is called *Beyond Atkins: How do the Prongs Perform during Sentencing?* You will receive course credit for your participation as part of the PY 101 course.

Why is this study important? We hope to find out how certain people are viewed during sentencing hearings. This information is important to people involved in these types of cases.

Why have I been asked to participate in this study? All students in the PY 101 had access to the description of the study. You have been asked to participate in this study because you met the requirements and signed up for the study.

How many other people will be asked to participate? Up to 480 other people will participate in this study.

What will I be asked to do in this study? You will be asked to listen to an audio recording of trial testimony and then provide responses to that recording. You will discuss the information with your peers and then unanimously decide on a sentence for the defendant. You will also complete several surveys that relate to your attitudes and beliefs about crime and justice. You will also be asked to provide basic information about yourself, for example, your age, gender, and political affiliation. The discussions with your peers may be audio-recorded by the researcher.

How long will this study last? Participation can last up to three hours. During the study, if you desire a break – please ask for one and we will take a break. The only cost to you from this study is your time. You will be awarded partial course credit for your participation. Regardless of the time to complete the study, you will be awarded 4.5 credits if you complete the whole study. If you stop your participation during the study, you will be awarded credits for the time you spent. For example, if you participate for two hours, you will be awarded two credits.

What are the risks and benefits of participation? The only risk to you is that you may get tired during the study. Because of this, breaks will be taken if requested. Although there are no direct benefits to you, you may learn more about how the legal system works and/or about sentencing. You may also feel good about knowing that this information could shape the legal system in the future. Individuals involved in the legal system may also benefit from the information learned in this study.

How will my privacy be protected? All of the surveys in this study will be completed on your own. Group discussions in this study may be audio recorded. The only people with access to these recordings will be the trained members of the research team. The recordings will be destroyed once they are typed in written format by one of the researchers. Because identifying information is not collected with your data (see below), there will be no way to match what you said with your identifying information. Your comments will be matched to your participant number, but not your name.

How will my confidentiality be protected? No identifying information will be paired with your data. You will provide your name and CWID at the end of the study for the purpose of awarding credit. This information will not be paired with your data in any way. Informed consent documents will be kept separate from the data. Also, only trained individuals will have access to the data from this project.

What are the alternatives to being in this study? The alternative to being in this study is to participate in another assignment for course credit. Please contact your instructor for information about this option.

UA IRB Approved Document
Approval date: 10-23 14
Expiration date: 10-22-15

What are my rights as a participant in this study? Taking part in this study is voluntary. It is your free choice. You can refuse to be in it at all. If you start the study, you can stop at any time. There will be no effect on your relations with the University of Alabama.

The University of Alabama Institutional Review Board ("the IRB") is the committee that protects the rights of people in research studies. The IRB may review study records from time to time to be sure that people in research studies are being treated fairly and that the study is being carried out as planned.

Who do I call if I have questions or problems?
If you have questions, concerns, or complaints about the study right now, please ask them. If you have questions, concerns, or complaints about the study later on, please email Mary E. Wood at mschnorf@crimson.ua.edu or Karen L. Salekin at ksalekin@ua.edu

If you have questions about your rights as a person in a research study, call Ms. Tanta Myles, the Research Compliance Officer of the University, at 205-348-8461 or toll-free at 1-877-820-3066.

You may also ask questions, make suggestions, or file complaints and concerns through the IRB Outreach website at http://osp.ua.edu/site/PRCO_Welcome.html or email the Research Compliance office at participantoutreach@bama.ua.edu.

After you participate, you are encouraged to complete the survey for research participants that is online at the outreach website or you may ask the investigator for a copy of it and mail it to the University Office for Research Compliance, Box 870127, 358 Rose Administration Building, Tuscaloosa, AL 35487-0127.

I have read this information sheet. I have had a chance to ask questions. I agree to take part in it.

I will receive a copy of this consent form to keep.

UA IRB Approved Document
Approval date: 10-23-14
Expiration date: 10-22-15

Debriefing Form

Dear Participant,

Thank you for your participation in this study. The purpose of this study was to investigate how certain factors are perceived by jurors during capital mitigation. The results of this study may influence how attorneys and/or mental health professionals approach capital mitigation in the future. We were particularly interested in how information about intellectual disability (ID) was evaluated.

Please do not share with others about your participation in this study. It is important that our data is not influenced by bias in any way and this includes information shared among potential participants. If you have additional questions about this project or the purpose of the study, please contact the principal investigator directly. Her contact information is below.

You are entitled to withdraw your data from the current study given that you were not fully informed of the purpose of the study at the outset. If you wish to have your data withdrawn, please contact the principal investigator at the email address below. You will not be penalized if you decide to have your data withdrawn.

Thanks,

Mary E. Wood, M.A.
Principal Investigator
mschnorf@crimson.ua.edu