Effects of Behavioral Genetic Evidence on Perceptions of Criminal Responsibility and Appropriate Punishment

Paul S. Appelbaum
Columbia University

Nicholas Scurich
University of California, Irvine

Raymond Raad
Columbia University

Demonstrations of a link between genetic variants and criminal behavior have stimulated increasing use of genetic evidence to reduce perceptions of defendants’ responsibility for criminal behavior and to mitigate punishment. However, because only limited data exist regarding the impact of such evidence on decision makers and the public at large, we recruited a representative sample of the U.S. adult population (N = 960) for a Web-based survey. Participants were presented with descriptions of 3 legal cases and were asked to determine the length of incarceration for a convicted murderer, adjudicate an insanity defense, and decide whether a defendant should receive the death penalty. A fully crossed, between-participants factorial design was used, varying the type of evidence (none, genetic, neuroimaging, both), heinousness of the crime, and past criminal record, with sentence or verdict as the primary outcome. Also assessed were participants’ apprehension of the defendant, belief in free will, political ideology, and genetic knowledge. Across all 3 cases, genetic evidence had no significant effects on outcomes. Neuroimaging data showed an inconsistent effect in 1 of the 2 cases in which it was introduced. In contrast, heinousness of the crime and past criminal record were strongly related to participants’ decisions. Moreover, participants’ beliefs about the controllability of criminal behavior and their political orientations were significantly associated with their choices. Our findings suggest that neither hopes that genetic evidence will modify judgments of culpability and punishment nor fears about the impact of genetic evidence on decision makers are likely to come to fruition.

Keywords: genetic evidence, mitigation, insanity defense, neuroimaging evidence, sentencing

Supplemental materials: http://dx.doi.org/10.1037/law0000039.supp

A variety of data supports the likely influence of genetic variables on criminal and other antisocial behaviors. Sibling and twin studies have provided evidence of familial aggregation of criminal activity, with Rhee and Waldman’s (2002) much-cited meta-analysis of 51 studies showing that additive genetic influences account for 32% of the variance and interactions among genes for another 9%. Two large twin studies estimated heritabilities of 37%–57% for five kinds of aggressive behaviors (Yeh, Coccaro, & Jacobson, 2010) and 67% for antisocial behavior (Tuvblad, Narusyte, Grann, Sarnecki, & Lichtenstein, 2011). Work is ongoing to identify genes that may be linked to antisocial behavior, including monoamine oxidase A (MAOA), catechol-O-methyltransferase, dopamine transporter DAT1, dopamine receptors DRD2 and DRD4, and serotonin transporter 5-HTTLPR (Fergusson & Beaver, 2009; Iofrida, Palumbo, & Pellegrini, 2014), although many of the findings remain a subject of contention (Vassos, Collier, & Fazel, 2014). It seems likely that data regarding genes associated with increased risk for criminal behavior will continue to appear (Tiihonen et al., 2014).

However, particular attention has been paid to MAOA, an enzyme that degrades monoamine neurotransmitters. After suggestive evidence of a link between absent MAOA activity and criminal behavior (Brunner, Nelen, Breakefield, Ropers, & van Oost, 1993), Caspi et al. (2002) took advantage of a longitudinal epidemiologic study of a birth cohort in Dunedin, New Zealand, to explore the phenomenon further. Examining high- and low-activity polymorphisms in the promoter region of the gene on the X chromosome in male subjects, they found no effect of MAOA genotype per se but did find evidence for an interaction between a history of childhood maltreatment and MAOA status—a gene–environment interaction. Subjects with an allele associated with reduced MAOA production who had a history of childhood mal-


THE EFFECTS THAT BEHAVIORAL GENETIC DATA WILL HAVE ON CRIMINAL PROCEEDINGS DEPEND ON THE ATTITUDES AND BELIEFS OF THE PARTIES INVOLVED IN CRIMINAL COURT PROCEEDINGS (JUDGES, JURIES, AND OTHERS) AND, ULTIMATELY, ON PUBLIC BELIEFS. NOT ONLY DOES THE PUBLIC CONSTITUTE THE POOL FROM WHICH JURORS ARE DRAWN, BUT EVEN FOR THOSE DETERMINATIONS THAT ARE NOT MADE BY JURORS IN MOST STATES (E.G., SENTENCING; KING & NOBLE, 2004), PUBLIC VIEWS MAY GUIDE THE DEVELOPMENT OF RULES FOR THE APPLICATION OF BEHAVIORAL GENETIC DATA. HOWEVER, PRIOR TO OUR PILOT STUDY (DESCRIBED IN THE FOLLOWING PARAGRAPH), THERE WERE NO EMPIRICAL DATA ON PUBLIC PERCEPTIONS OF BEHAVIORAL GENETICS AND ITS IMPACT ON CRIMINAL RESPONSIBILITY. ONE STUDY FOCUSED ON POTENTIAL JURORS’ COMPREHENSION OF GENETIC INFORMATION AND WAS ENCOURAGING IN THAT REGARD (HANS, KAYE, DANN, FARLEY, & ALBERTSON, 2011). ANOTHER STUDY LOOKED AT STATE TRIAL COURT JUDGES AND FOUND THAT AFTER HEARING A DESCRIPTION OF A CASE OF AGGRAVATED BATTERY BY A PSYCHOPATHIC DEFENDANT, THOSE JUDGES WHO WERE EXPOSED TO GENETIC AND OTHER BIOLOGICAL DATA TO SUPPORT THE DIAGNOSIS OF PSYCHOPATHY INDICATED THAT THEY WOULD IMPOSE LESS SEVERE SENTENCES ON THE PERPETRATOR. THE DIFFERENCE WAS SIGNIFICANT; HOWEVER, THE ABSOLUTE MAGNITUDE OF THE EFFECT—ONE LESS YEAR IN PRISON, ON AVERAGE—WAS SMALL (12.83 YEARS [SEM = 1.104] VS. 13.93 YEARS [SEM = 1.039]; ASPINWALL, BROWN, & TABERY, 2012), AND THE METHODS AND CONCLUSIONS OF THE STUDY HAVE NOT GONE UNCHALLENGED (DENNO, 2013).


THERE HAVE ALSO BEEN EFFORTS TO STUDY THE IMPACT OF OTHER KINDS OF NEUROSCIENTIFIC INFORMATION, AND THEY HAVE LED TO CONFLICTING RESULTS. ONE WIDELY CITED STUDY SUGGESTED THAT NEUROSCIENTIFIC INFORMATION, EVEN WHEN IRRELEVANT TO THE ISSUE IN QUESTION, LENT CREDIBILITY TO POOR EXPLANATIONS OF BEHAVIORAL PHENOMENA AMONG NONEXPERTS (WEISBERG, KEIL, GOODSTEIN, RAWSON, & GRAY, 2008). USING A SOMEWHAT DIFFERENT PARADIGM, ANOTHER RESEARCH GROUP SHOWED THAT BRAIN IMAGES, BUT NOT OTHER SORTS OF ILLUSTRATIONS, INCREASED PERCEPTIONS OF THE QUALITY OF SCIENTIFIC REASONING IN ARTICLES ABOUT SCIENTIFIC ADVANCES (MCCABE & CASTEL, 2008).
recently, however, several studies with more precise methodologies have failed to replicate those findings (Farah & Hook, 2013; Gruber & Dickerson, 2012; Michael, Newman, Vuorre, Cumming, & Garry, 2013). Two studies of mock jurors being asked to make decisions about mens rea (culpable mental states) and sentencing have shown no incremental effect of brain images beyond the impact of verbal testimony regarding neuropsychological impairment (Greene & Cahill, 2012; Schweitzer et al., 2011). However, a more recent study of sentencing in capital trials found that neuroimaging reduced perceptions of responsibility and sentences of death for defendants diagnosed with psychopathy but increased perceived responsibility in defendants with schizophrenia (Saks, Schweitzer, Aharoni, & Kiehl, 2014). Hence, the likely effect of neuroscientific evidence in legal settings is still unclear.

The Present Study

The growing attention to genetic evidence in the legal system calls for a more sophisticated examination of its potential impact. In the present article, we report the results of a large-scale study that used a representative sample of the U.S. population to look at the effect of behavioral genetic information. Participants were exposed to three different legal cases in which we varied the ways such evidence may be used, the types of scientific evidence presented, and the characteristics of the offender and of the crime.

Two of the cases in this study deal with sentencing contexts (capital and noncapital), given suggestions that genetic evidence is most likely to be used at that phase of trial for purposes of mitigation (Denno, 2011). The third case is an insanity defense, with the goal of exploring the potential impact of genetic data in a traditional justification defense, including their use to support diagnostic claims, which appears to be common (Farahany, 2011). Because consideration of genetic evidence is likely to be integrated by decision makers into a larger “story,” including other characteristics of the defendant and the crime (Pennington & Hastie, 1986), we examined the interactions of such characteristics with the scientific evidence. In particular, because the impact of evidence presented in mitigation or justification of an offense may be muted by the heinousness of the offense and the defendant’s previous criminal record, we systematically varied these characteristics. (There are reasons to believe that a history of prior violence, in particular, may be likely to affect sentencing decisions, given that it may be perceived as a proxy for future dangerousness [Blume, Johnson, & Sundby, 2008; Garvey, 1998].) Finally, given that evolving practice appears to include presentation of both genetic and neuroimaging data (Bernet et al., 2007), we explored the impact of these types of evidence separately and jointly. To highlight the scientific evidence for our participants, we used several types of images: neuroimages showing impairment in the frontal lobes, which may be associated with increased impulsivity; images of genetic test results (i.e., a Manhattan plot and a figure showing a deletion in one of the defendant’s chromosomes), which we explained as indicating the presence of variants associated with heightened impulsivity or a psychiatric disorder, respectively; and a genogram, which graphically displays family relationships and the penetrance of the condition in question.

As potential correlates of participants’ decisions, in addition to their demographic characteristics, we used measures of knowledge and attitudes that might explain their choices: measures of belief in free will, political ideology, and genetic knowledge. Although the literature on free will is complex, data suggest that greater belief in free will predicts more punitive approaches to sentencing (Shariff et al., 2014). Similarly, conservative political views, at least at the aggregate level, appear to be associated with longer prison sentences (Bowers, 1998) and greater use of the death penalty (McCann, 2008). Lastly, we hypothesized that greater familiarity with genetic concepts would make respondents more likely to see genetic testimony as mitigating or justifying criminal conduct.

Method

Participants

Identification and recruitment of participants was conducted by YouGov (http://yougov.com), a survey research company that maintains a Web-based panel of respondents. YouGov constructed a sample representative of the adult U.S. population with a two-stage sampling design. First, a sampling frame was constructed from the American Community Survey (U.S. Census Bureau, 2010), with additional data from the Current Population Survey voter supplement (U.S. Census Bureau, 2008) and the Pew Religious Life Study (Pew Research, 2007). From the frame, a stratified random sample was drawn similar in size to the desired study sample. At the second stage, the sampling algorithm behind the proprietary sampling system searched the opt-in panel (i.e., respondents to a generic invitation to participate in a survey) for participants who most closely matched the individuals in the randomly drawn target sample. The algorithm invited 2–3 matches for every respondent in the target frame. The matching criteria included age, race, gender, and education. The final sample (N = 960) had the characteristics of the adult U.S. population.

Sample size was determined by an a priori power analysis. Assuming a medium effect size (delta) of 0.75 (Cohen, 1988; delta, sometimes referred to as d, is the difference between the largest mean and the smallest mean, in units of the within-cell standard deviation) and Type I error of 5%, each case required more than 30 participants per condition to obtain power greater than .80. On the basis of these assumptions and given a maximum of 24 cells per case, the minimum required sample size was 720; however, we opted to have 40 participants per cell, for a total of 960 participants. The demographic characteristics of the final sample are provided in Table 1. In addition, participants were asked to respond to an 18-item questionnaire testing their knowledge of basic genetics; mean and median percentages correct were both 83.0% (SD = 10.9%), indicating a good level of genetic knowledge among most participants. Participants also completed the free will subscale of the FAD-Plus scale (Cronbach’s α = .866), which measures beliefs about free will and related concepts (Paulhus & Carey, 2011). Higher scores indicate greater beliefs in free will.

YouGov provided data on respondents’ self-described political orientation (from very liberal to very conservative) on the basis of a standard question they ask of their panel members.

Drafts of the case vignettes described in the following were reviewed for the realism of their portrayals of legal and mental health issues and for clarity by a panel of reviewers that included five attorneys (four of them legal academics with relevant expertise), an experienced academic forensic psychiatrist, and an academic psychologist with expertise in issues related to evidence
Table 1
Characteristics of Participants

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>462</td>
<td>48.1</td>
</tr>
<tr>
<td>Female</td>
<td>498</td>
<td>51.9</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>694</td>
<td>72.3</td>
</tr>
<tr>
<td>Black</td>
<td>103</td>
<td>10.7</td>
</tr>
<tr>
<td>Hispanic</td>
<td>97</td>
<td>10.1</td>
</tr>
<tr>
<td>Asian</td>
<td>17</td>
<td>2.0</td>
</tr>
<tr>
<td>Native American</td>
<td>11</td>
<td>1.1</td>
</tr>
<tr>
<td>Mixed</td>
<td>23</td>
<td>2.4</td>
</tr>
<tr>
<td>Other</td>
<td>15</td>
<td>1.6</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;High school graduate</td>
<td>38</td>
<td>4.0</td>
</tr>
<tr>
<td>High school graduate</td>
<td>382</td>
<td>39.8</td>
</tr>
<tr>
<td>Some college</td>
<td>222</td>
<td>23.1</td>
</tr>
<tr>
<td>Two-year college degree</td>
<td>72</td>
<td>7.5</td>
</tr>
<tr>
<td>Four-year college degree</td>
<td>157</td>
<td>16.4</td>
</tr>
<tr>
<td>Postcollege education</td>
<td>89</td>
<td>9.3</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>458</td>
<td>47.7</td>
</tr>
<tr>
<td>Domestic partnership</td>
<td>42</td>
<td>4.4</td>
</tr>
<tr>
<td>Separated</td>
<td>13</td>
<td>1.4</td>
</tr>
<tr>
<td>Divorced</td>
<td>94</td>
<td>9.8</td>
</tr>
<tr>
<td>Widowed</td>
<td>46</td>
<td>4.8</td>
</tr>
<tr>
<td>Never married</td>
<td>307</td>
<td>32.0</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time</td>
<td>342</td>
<td>35.6</td>
</tr>
<tr>
<td>Part-time</td>
<td>99</td>
<td>10.3</td>
</tr>
<tr>
<td>Currently unemployed</td>
<td>119</td>
<td>12.4</td>
</tr>
<tr>
<td>Retired</td>
<td>164</td>
<td>17.1</td>
</tr>
<tr>
<td>Permanently disabled</td>
<td>70</td>
<td>7.5</td>
</tr>
<tr>
<td>Homemaker</td>
<td>72</td>
<td>7.5</td>
</tr>
<tr>
<td>Student</td>
<td>67</td>
<td>7.0</td>
</tr>
<tr>
<td>Other</td>
<td>27</td>
<td>2.8</td>
</tr>
<tr>
<td>Political orientation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very liberal or liberal</td>
<td>244</td>
<td>25.4</td>
</tr>
<tr>
<td>Moderate</td>
<td>280</td>
<td>29.2</td>
</tr>
<tr>
<td>Very conservative or conservative</td>
<td>335</td>
<td>34.9</td>
</tr>
<tr>
<td>Not sure</td>
<td>101</td>
<td>10.5</td>
</tr>
</tbody>
</table>

Note. N = 960.

The final versions of the case vignettes were modified per the reviewers’ comments.

**General Procedure**

In this online study, participants were presented with descriptions of three different legal cases and asked to render a decision. One case asked participants to determine the length of incarceration for a convicted murderer (Case 1), the second case had participants adjudicate whether a defendant should be found not guilty by reason of insanity (Case 2), and the final case asked whether a defendant convicted of capital murder should receive the death penalty or life in prison without the possibility of parole (Case 3). The approximate length of the stimulus for each case was 700 words. Whenever genetic or neuroimaging findings were introduced, they were accompanied by an image to enhance the likelihood that participants would attend to the key manipulations. The order in which the three cases were presented was randomized, and participants were randomly assigned to one experimental condition within each case. After rendering a decision in each case, participants answered eight questions about their reactions to the defendant. For example, participants were asked, “To what extent is [the defendant] an immoral person?” and “How dangerous is [the defendant] to the public?” All ratings were made on a nine-point Likert scale. Finally, participants provided general background and demographic information.

**Case 1**

**Design**

The experimental design was a 4 (explanation of behavior: impulsive, genetic, neuroimaging [neuro], both) × 2 (heinousness of crime: low, high) × 3 (criminal history: none, nonviolent, violent) fully crossed, between-participants factorial design. Participants read about a fracas that occurred outside a bar when the defendant provoked a fight because the victim had flirted with his girlfriend. During the course of the altercation, the defendant stabbed the victim, who was pronounced dead on the scene by the emergency response team. The defendant unsuccessfully claimed that the stabbing was done in self-defense. Participants were told that the defendant had been convicted of murder and was now awaiting sentencing.

There were three experimental manipulations in this case. First, the defense’s proffered explanation of the murder was varied. Participants were randomly assigned to vignettes in which the defense attorney claimed that it was simply an impulsive act (impulsive), a psychiatrist testified that the defendant had a rare gene that predisposed him to impulsive and violent behavior (genetic explanation), a psychiatrist testified that the results of brain imaging revealed that the defendant had abnormal activity in the front part of the brain that predisposed him to impulsive and violent behavior (neuro explanation), or a psychiatrist testified about both the genetic and neuro explanations (both). The genetic and neuro explanations were accompanied by images of genetic test results and MRI scans, respectively.

Two additional variables in the vignettes that were hypothesized to have an effect on participants’ sentencing decisions were also manipulated. The heinousness of the murder was varied so that there was either one stab wound that killed the victim instantly (low) or 17 stab wounds all over the victim’s body, including one in the eye (high). Second, in his remarks at the sentencing hearing, the prosecutor noted the defendant’s history of criminal conduct: a previous conviction for shoplifting (nonviolent), a previous conviction for assault (violent), or no previous convictions (none). In addition, the prosecutor argued that the defendant’s actions proved that he is a “bad actor” who should be punished for his behavior regardless of the explanations offered.

After reading the materials, participants were asked to indicate the appropriate prison sentence for the defendant on a 1–60-year scale. Participants then completed the eight reaction questions.

**Results**

The median prison sentence selected by our respondents was 24 years, with an interquartile range (IQR) of 27. A histogram of participants’ responses to the prison sentence question (see Figure 1) demonstrates substantial variation in the length of confinement chosen.
A three-way analysis of variance (ANOVA) was conducted to examine whether prison sentences varied as a function of the experimental manipulations. The ANOVA detected main effects for heinousness of crime, $F(1, 960) = 37.23, p < .001, \eta^2_p = .038$, and criminal history, $F(2, 960) = 3.56, p < .05, \eta^2_p = .008$. The main effect for explanation of behavior (i.e., impulsive, genetic, neuro, or both) was not significant, $F(3, 960) < 1$, nor were any of the interactions (all $p > .10$). The heinousness of the crime affected prison sentences such that the high-heinousness killing resulted in significantly longer sentences ($M = 29.82, 95\%$ confidence interval $[28.2, 31.4]$) than the low-heinousness killing ($M = 22.79, 95\%$ CI $[21.4, 24.5]$). Compared with no criminal history, both nonviolent and violent criminal histories resulted in longer prison sentences ($Ms = 24.33, 27.07$, and $27.75$, respectively). A post hoc Bonferroni test indicated that the only significant contrast was between no criminal history and violent criminal history ($p < .05$); the difference between no criminal history and a nonviolent criminal history, as well as that between nonviolent and violent criminal histories, was not significant ($ps > .10$).

Several of the participant variables correlated with the length of prison sentences. Belief in free will, as measured by the FAD-Plus, had a correlation of $r = .237 (p < .001)$ with the sentence, indicating that the length of confinement imposed increased as participants’ belief in free will increased. Genetic knowledge was negatively, but weakly, correlated ($r = -.082, p < .01$) with prison sentences, with greater knowledge about genetics associated with shorter prison sentences. Political ideology was positively, albeit weakly, correlated with prison sentences ($r = .095, p < .001$), with higher levels of conservatism related to longer prison sentences.

Combining the responses to the eight reaction questions into a single composite score resulted in a Cronbach’s alpha of .761 indicating the presence of a latent construct (see DeVellis, 2003).

We refer to this construct as apprehension of defendant, with higher scores equaling greater apprehension of the defendant. A scale reliability analysis indicated that removing any particular item would not significantly increase alpha. The Pearson correlation of apprehension of defendant and the prison sentence was $.615 (p < .001).

A three-way ANOVA with heinousness of crime, explanation of behavior, and criminal history as the independent variables and apprehension of defendant as the dependent variable detected a significant main effect only for heinousness of crime, $F(1, 960) = 22.79, p < .001, \eta^2_p = .024$. The main effects for explanation of behavior, criminal history, and all the interactions were not statistically significant (all $ps > .10$).

### Discussion

There was substantial variation in the sentences selected by our participants in response to this vignette describing an unpunished murder, but the explanation of the defendant’s behavior that was offered by the defense had no significant effect on their choices. However, although respondents were not swayed by the scientific evidence of genetic or brain-related influences on the defendant’s behavior, they were affected by more traditional types of evidence. Specifically, heinousness of the crime and past criminal history were significant predictors of the length of sentence, and heinousness also predicted the degree of apprehension of the defendant manifested by respondents. Despite the concern sometimes expressed that genetic and neuroimaging evidence might be a “double-edged sword” (Aspinwall et al., 2012) that could be turned against the defendant who introduces them, we found neither a positive nor a negative effect of such evidence. Although jurors in most states are not asked to make sentencing determinations, these findings suggest that the general public is not inclined to see genetic or neuroimaging evidence that the defendant is at heightened risk of impulsive behavior as warranting modification of sentences.

### Case 2

#### Design

The experimental design was a 2 (heinousness of crime: low, high) × 2 (type of genetic evidence: lab, family tree) × 3 (purpose of test: diagnosis, impulsivity) × 3 (history of mental illness: none, mild, severe) fully crossed, between-participants factorial design. Participants read about a case in which a 32-year-old man made several affectionate advances toward a female coworker. All the advances were rejected. The final rejection sent the defendant into a spiral of uncharacteristic behavior, including extreme anxiety, sleep disturbance, disorganized speech, and profligacy with money. One night, the defendant spotted the coworker at a local gas station. He approached and tried to hug her, but she pushed him away. He grabbed her by the hair, dragged her into his truck, and drove off. The entire event was captured by surveillance cameras and viewed by three eyewitnesses. The victim’s body was discovered 2 days later dumped by the side of a road. The defendant was arrested and charged with kidnapping and murder, to which he pled not guilty by reason of insanity (NGRI). A psychiatrist called by the defense testified that the defendant met the
criteria for NGRI under a modified version of the American Law Institute’s formulation for insanity (i.e., a cognitive or volitional impairment claim).

There were four experimental manipulations in this study. First, the heinousness of the crime was manipulated: the victim was either shot once in the head (low) or shot once each in the head and the genitals (high). Second, the psychiatrist conducted one of two types of genetic investigations, either a laboratory test for genetic abnormalities (lab test) or an analysis of the defendant’s family tree (family tree); surveys of reported cases suggest that genetic evidence is introduced in both forms in criminal trials (Denno, 2011, 2013). In both conditions, an image of the genetic evidence was presented. Third, the purpose of the evidence was varied: the evidence was either introduced to buttress a diagnosis of bipolar disorder (diagnosis) or to explain directly the defendant’s impulsive behavior (impulsivity). Finally, the defendant’s history of mental illness was described as consisting of no prior episode (none), a similar but less severe episode (mild), or a very similar episode (severe).

After reading the case materials, participants were given judicial instructions regarding the test of insanity to be applied. They were also told that if the defendant was found NGRI, he would be sent to a maximum-security psychiatric hospital. Participants then indicated whether they would find the defendant guilty of murder or NGRI. On a separate screen, participants who chose a guilty verdict were asked for how many years the defendant should be confined to the hospital, both on a 1–60-year scale. Finally, participants answered the eight reaction questions.

Results

A majority of participants (62%; n = 593) found the defendant guilty. With regard to the prison sentence if the defendant was convicted, the median response was 45 years (IQR = 30, range = 1–60), whereas the median length of hospital incarceration if the defendant was found NGRI was 20 years (IQR = 35, range = 1–60).

A binary logistic regression with type of evidence, purpose of evidence, heinousness of crime, and history of mental illness as the independent variables and verdict (i.e., guilty or NGRI) as the dependent variable was not statistically significant, χ²(7, N = 960) = 9.37, p = .23. This indicates that the experimental manipulations cannot explain variance in the participants’ verdicts. In other words, the manipulations did not affect the likelihood that participants would find the defendant NGRI.

Participants who found the defendant guilty compared with those who found him NGRI had stronger beliefs in free will (M = 4.17, 95% CI [4.11, 4.22] vs. M = 3.58, 95% CI [3.51, 3.66]), t(958) = 12.37, p < .001, and more conservative political ideology (M = 3.63, 95% CI [3.52, 3.75] vs. M = 3.17, 95% CI [3.02, 3.31]), t(958) = 5.01, p < .001, but lower levels of genetic knowledge (M = 0.82, 95% CI [0.81, 0.83] vs. M = 0.84, 95% CI [0.83, 0.85]), t(958) = −3.16, p = .002.

Combining the responses to the eight reaction questions into a single composite score resulted in a Cronbach’s alpha of .789. A scale reliability analysis indicated that removing any particular item would not significantly increase alpha. Participants who found the defendant guilty had significantly greater apprehension of the defendant (M = 6.26, 95% CI [6.17, 6.34]) than participants who chose a NGRI verdict (M = 4.52, 95% CI [4.41, 4.63]), t(958) = 23.89, p < .001. In addition, an ANOVA with the experimental manipulations as the independent variables and apprehension of defendant as the dependent variable detected a significant main effect for purpose of test, F(1, 960) = 3.78, p < .05, η² = .004. In general, when the evidence was introduced to buttress a diagnosis, the apprehension of defendant scores were lower (M = 5.5) compared with when the evidence was introduced to explain impulsivity (M = 5.7). A two-way Heinousness of Crime × Type of Genetic Evidence (i.e., lab vs. family history) interaction was also detected, F(1, 960) = 5.14, p < .05, η² = .005 (see Figure 2). None of the other main effects or interactions was significant (all ps > .10).

As is apparent, the type of genetic evidence had an effect, but only when the crime was low in heinousness. With low heinousness, the laboratory test engendered more apprehension of the defendant than the family tree. However, when the crime was high in heinousness, the type of genetic evidence had no effect. Note that the only difference between the heinousness conditions was an extra gunshot wound to the victim’s genital area.

Discussion

Empirical data on the use of the insanity defense indicate that most cases that go to trial end with a guilty verdict (74% in one large study; Silver, Cirincione, & Steadman, 1994), which was the outcome for the majority of respondents here. That outcome was not affected by the introduction of genetic evidence either to support the diagnosis of bipolar disorder or to speak directly to the defendant’s ability to control his behavior. Nor did it matter whether the testimony about the defendant’s genetic predisposition was based on genetic testing conducted in a laboratory or on his family history. In contrast to the previous case, the degree of heinousness of the offense had no impact on the verdict, perhaps because some respondents took it as evidence that the crime was driven by mental illness, whereas others saw it as indicating vengefulness. Neither did heinousness directly affect respondents’ apprehension of the defendant, although it did show an effect in modifying the impact of the type of genetic evidence on apprehension of the defendant. Once more, there was no evidence of a
paradoxical effect of genetic evidence; that is, its introduction did not result in higher rates of guilty verdicts—although evidence of a genetic basis for increased impulsivity did lead to heightened apprehension of the defendant.

Case 3

Design

The experimental design was a $2 \times 3$ (heinousness of crime: low, high) $\times 3$ (criminal history: none, nonviolent, violent) factorial design. Participants were told that they would be reading about a case in which they would be asked to consider the death penalty for the defendant. Consistent with federal guidelines for death qualification, participants were asked, “Do you have a conviction against the death penalty so strong that you could not take an oath to fairly try a death penalty case and follow the law?” and required to respond “yes” or “no.” They then read a scenario in which three individuals robbed a bank at gunpoint. Responding to a silent alarm, a police officer entered the bank and demanded that the suspects drop their weapons. The defendant fatally shot the police officer from behind. The trio was apprehended as they fled, and the defendant was subsequently convicted of first-degree murder. Participants were asked to determine whether the defendant should receive the death penalty or there were mitigating factors that called for leniency, in which case he would be sentenced to life in prison without parole (LWOP).

There were three manipulations in this study. First, the heinousness of the crime was manipulated: The victim was either shot once and then fell to the ground and died (low) or he was shot once and fell to the ground, at which time the defendant stood over him and fired until his weapon was empty (high). Second, a presenting report indicated that the defendant’s criminal history included no previous convictions (none), convictions for disorderly conduct and graffiti (nonviolent), or convictions for assault and rape (violent). Finally, the proffered explanation and putative mitigation for the defendant’s conduct that was offered by his attorney was varied: The attorney claimed it was simply an impulsive act (impulsive explanation); a psychiatrist testified that the defendant had a rare gene that “led the defendant to act on impulses without thinking them through, and this is certainly a result of his biological makeup” (genetic explanation); a psychiatrist offered both the genetic and neuro explanations (both).

After reading the case materials, participants were given sentencing instructions and asked to determine a sentence, which was either LWOP or the death penalty. Finally, participants answered the eight reaction questions.

Results

A large majority of participants (80%; $n = 763$) indicated that they did not have a principled objection to the use of the death penalty that would affect their decision making in the case. In other words, 80% of the participants were qualified to adjudicate a death penalty case; such jurors are commonly referred to as death-qualified jurors. Because non-death-qualified jurors who are identified during voir dire are typically excluded from death penalty hearings (Witherspoon v. Illinois, 1968), only death-qualified participants are included in the analyses reported here.

Imposition of the death penalty was chosen by 58% ($n = 448$) of participants. A binary logistic regression was conducted with heinousness of crime, criminal history, and explanation of behavior as the independent variables and sentence (i.e., LWOP or death) as the dependent variable. The model was significant, $\chi^2(12, N = 763) = 41.03, p < .001$. There was a significant main effect for heinousness of crime, with participants in the high-heinousness condition being over 2.5 times more likely to impose death, $exp(B) = 2.52, 95\% CI [1.65, 2.86], Wald = 18.07, df = 1, p < .001$. There was also a main effect for criminal history (Wald = 9.31, $df = 1, p < .01$), with participants in the nonviolent and violent conditions being nearly two times more likely to impose the death penalty, $exp(B) = 1.85, 95\% CI [1.17, 2.91], Wald = 6.90, df = 1, p < .01$, than participants in the none condition, $exp(B) = 1.79, 95\% CI [1.15, 2.72]$, Wald = 6.64, $df = 1, p < .01$. There was also a main effect for explanation of behavior (Wald = 8.53, $df = 1, p < .05$); the only significant effect pertained to the neuro explanation, with participants in this condition being less likely to impose death than participants in the impulsive condition, $exp(B) = 0.56, 95\% CI [0.35, 0.87], Wald = 6.47, df = 1, p < .05$.

Combining the responses to the eight reaction questions into a single composite score resulted in a Cronbach’s alpha of .769. A scale reliability analysis indicated that removing any particular item would not significantly increase alpha. Participants who voted to impose the death penalty had significantly greater apprehension of the defendant ($M = 7.72, 95\% CI [6.54, 6.72]$) than participants who chose LWOP ($M = 6.66, 95\% CI [6.54, 6.77]$), $t(761) = 13.66, p < .001$, as well as greater belief in free will ($M = 4.15, 95\% CI [4.09, 4.22]$ vs. $M = 3.84, 95\% CI [3.77, 3.92]$), $t(761) = 16.25, p < .001$. They also showed higher levels of conservatism ($M = 3.80, 95\% CI [3.68, 3.93]$ vs. $M = 3.32, 95\% CI [3.17, 3.47]$), $t(761) = 4.82, p < .001$. There were no differences in genetic knowledge between participants who imposed the death penalty and those who voted for LWOP, $t(761) = 0.355, p = .723$.

An ANOVA with the experimental manipulations as the independent variables and apprehension of defendant as the dependent variable detected significant main effects for heinousness of crime, $F(1, 763) = 7.28, p < .01$, $\eta^2_p = .01$, and explanation of behavior, $F(1, 763) = 6.50, p < .001$, $\eta^2_p = .037$, and a two-way Heinousness of Crime $\times$ Explanation of Behavior interaction, $F(1, 763) = 2.63, p < .05$, $\eta^2_p = .011$; see Figure 3). None of the other main effects or interactions was significant (all $ps > .10$).

The interaction resulted from the impulsive explanation and high-heinousness condition, which engendered inordinate apprehension of the defendant. That is, there were no differences in apprehension ($ps > .10$ for all orthogonal contrasts) among the various explanations in the low-heinousness condition. Nor were there differences in the high-heinousness condition among the genetic, neuro, and both explanation conditions ($ps > .10$); the only difference was between the impulsive explanation and...
the other explanations ($p < .05$) in the high-heinousness condition.

**Discussion**

Death penalty hearings appear to be the context in which behavioral genetic data are most likely to be introduced in court (Denno, 2011), hence the results of this vignette have direct relevance to a common, high-stakes scenario. It is of considerable interest, therefore, that the results of introducing genetic evidence resemble those seen in the noncapital-sentencing context of Case 1. Genetic evidence did not affect the decision regarding imposition of the death penalty or LWOP, but, again, both the heinousness of the crime and past criminal history had significant effects. For the first time, however, we also identified an effect of neuroimaging evidence, albeit a limited one: Introduction of neuroimaging evidence led to a reduction in the likelihood of the death penalty compared with an argument about impulsivity in the absence of scientific evidence; however, the effect of neuroimaging evidence was not significantly different from the effect of genetic evidence alone or genetic and neuroimaging evidence combined. Neither type of scientific evidence enhanced apprehension of the defendant, although an argument that the defendant’s behavior represented impulsive (and by implication, unpredicated) behavior did increase apprehension when the crime was particularly heinous.

**General Discussion**

Given the growing literature about genetic influences on violent and other antisocial behaviors, and the increasing use of genetic evidence in court, this study examined the likely impact of such evidence on sentencing and insanity defense determinations by a representative sample of the U.S. adult population. Across all three cases, genetic evidence had no significant effects on the outcomes of the case—length of prison sentence, findings of insanity, or imposition of the death penalty. This finding is consistent with the results of our earlier, pilot study, which similarly found no impact of genetic evidence on determinations of a defendant’s degree of culpability or the sentence imposed in a noncapital case of homicide (Appelbaum & Scurich, 2014). In contrast, heinousness of the crime and past criminal record were strongly related to the participants’ decisions. Taken together, our data suggest that the impact of genetic evidence may have been considerably overestimated, especially when compared with characteristics of the crime and the offender’s history, at least given current public knowledge of and attitudes toward genetics.

Because we included conditions involving neuroimaging evidence in two of the cases, both alone and in combination with genetic evidence, we were able to compare the effect of another type of scientific evidence, one more commonly introduced in court than genetic data (Farahany, 2011). In addition, especially in death penalty cases, it appears increasingly common for genetic and neuroimaging data to be used in tandem (Bernet et al., 2007), which we explored in these cases as well. The previous literature shows varying and inconsistent effects of neuroimaging evidence in criminal cases (Greene & Cahill, 2012; Saks et al., 2014; Schweitzer et al., 2011), resembling our findings here. Neuroimaging evidence did not affect the outcome in the noncapital-sentencing case and had an inconsistent effect in the death penalty case, in which it reduced the likelihood of a death penalty in comparison with an impulsivity claim, but not when it was paired with genetic evidence. Here too, it appears that fears that neuroimages, with their graphic portrayal of brain function, would come to dominate jurors’ decisions were considerably overblown (Farah & Hook, 2013).

In addition to genetic evidence failing to alter the legal outcomes in these cases, it appeared to have little impact on participants’ perceptions of the defendants involved. This is in contrast to our previous study, in which genetic evidence of a propensity to impulsive and violent behavior increased fear of the defendant (Appelbaum & Scurich, 2014). In two of our three cases here, there was no significant effect of genetic evidence on respondents’ scores on the apprehension scale. The exception was the insanity defense case, which showed a complex pattern, with genetic data suggesting a propensity to behave impulsively inducing greater apprehension but not genetic data supporting the defendant’s diagnosis of bipolar disorder. In addition, in that case, genetic test data were more fear inducing than family history data. Thus, at most we can describe an inconsistent effect on perceptions of the defendant, suggesting that our sample of the general population did not hold the kind of beliefs in genetic determinism that are sometimes attributed to ordinary people, a finding consistent with other studies of popular views (Condit, 1999).

As expected, participants with higher levels of belief in free will and more conservative political views consistently chose more punitive options: longer prison sentences, fewer NGRI verdicts, and more death sentences. Higher levels of genetic knowledge, however, were associated with shorter sentences and more NGRI findings, although not with fewer death penalty verdicts. Thus, participants’ commitments to beliefs about the controllability of criminal behavior and their political orientations had a much greater impact on the outcomes of these cases than did either genetic or neuroimaging evidence.

The limitations of this study include those commonly associated with online surveys. The identity of the person completing the survey can never be known with certainty, and that person may not be the same as the person to whom the invitation was sent. In addition, in the absence of direct observation, the care with which the survey is completed is unknown. However, research suggests that Internet survey data tend to be of high quality, with good internal consistency and high test–retest reliability, yielding results.

**Figure 3.** Apprehension of the defendant as a function of the heinousness of the crime and the proffered explanation of his behavior. Error bars represent 95% confidence intervals.
similar to those obtained by traditional methods (Buhrmester, Kwang, & Gosling, 2011; Gosling, Vazire, Srivastava, & John, 2004). Moreover, the YouGov panel in this study provided a statistically representative sample of the general population—albeit not a true random sample—that would not otherwise have been easy to obtain.

With regard to methodological concerns, the use of written vignettes is quite different from the extended oral presentations that would take place at trial. Whether our participants would have responded differently to realistic trial testimony or even to written vignettes with somewhat different content (e.g., with the characteristics of the perpetrators, victims, or experts varied) is unknown; we cannot rule out the possibility that our manipulations had less impact than trial testimony or alternatively designed vignettes might have had. The concomitant use of genetic/neuroimaging findings and images makes it impossible for us to assess the distinct effects, if any, of either the findings or the images alone. Finally, we should note that although jurors are usually called on for NGRI and death sentence determinations, in most states (with six exceptions; Hoffman, 2003) jurors do not make noncapital sentencing decisions. Thus, with regard to Case 1 or the sentencing versus hospitalization decision in Case 2, sentencing or reviewing judges may have responded differently to the situation presented than did our respondents. However, respondents’ views provide an exploratory perspective on public attitudes regarding the question of what impact genetic and neuroimaging data should have on such decisions and arguably are a more sensitive measure of the impact of such evidence than dichotomous determinations of guilty versus NGRI verdicts or sentences of death versus life in prison.

What are the implications of our data for the use of psychiatric and behavioral genetic evidence in criminal trials? Assuming that our findings are replicated in future work, they suggest that the increase resort to genetic evidence by criminal defense lawyers (Denno, 2011) may be unlikely to yield the desired results. Our sample of the general population—from which juries are drawn—declined to rely on genetic (or, for the most part, neuroimaging) evidence in their determinations of culpability and sentencing. However, there was also no evidence of an adverse effect on defendants of genetic arguments (the so-called double-edged sword). Resort to psychiatric and behavioral genetic evidence seems to be most common in death penalty cases (Denno, 2011), perhaps in part because sentencing rules may limit the introduction of evidence demonstrating an impaired capacity to control behavior in some other cases (e.g., U.S. Sentencing Commission, 2014). Moreover, in capital cases, defense lawyers might correctly perceive that their clients have little to lose from putting their genetic propensities into evidence; the chance that some jurors may be influenced toward mitigation may be sufficient for the use of such evidence to continue. The present data, though, suggest that prospects may be dim for the successful extension of this practice into other areas of the criminal law.

Future research on the use of genetic and neuroscientific data in court might build on these findings by using methods that more closely resemble actual trials and jury deliberations. These could include live presentations of arguments by advocates for each side or videotaped testimony by real or simulated witnesses. Allowing groups of participants to deliberate jointly would also yield a more ecologically valid context for the examination of the impact of genetic information in the criminal courts. Although we used static images of genetic test results and neuroimages to highlight the scientific evidence being presented, there may be more effective ways of presenting the data. In addition, the effect of genetic data might be enhanced by quantifying its effect, which we did not do in this study; that is, the increase in the propensity for violent, antisocial, or impulsive behavior could be presented in numerical terms (e.g., “a doubling of the risk of violence”) to see if concretizing the effect in that way yields a different outcome. Finally, the next wave of our work will explore the use of genetic evidence in other adjudicatory contexts, including juvenile court settings and disciplinary hearings. Perhaps in settings in which criminal punishment is not the prime consideration, mitigating arguments that rely on genetic data will have greater effect. At present, though, our data are fairly consistent in suggesting that the hopes for data on genetic propensities to impulsive violence to modify judgments of culpability and appropriate punishment appear to be unfulfilled.

References


Mобильы, A., State, 455 S. E., 2d 61 (Ga. 1995).


Received November 7, 2014

Revision received January 5, 2015

Accepted January 7, 2015