Jurors’ perceptions of scientific testimony: The role of gender and testimony complexity in trials involving DNA evidence

Evelyn M. Maeder1*, Laura A. McManus2, Kendra J. McLaughlin2, Susan Yamamoto2 and Hannah Stewart1

Abstract: With continuous advancements in forensic science, expert testimony has become more common in criminal proceedings. This study (N = 170) sought to examine the combined influence of mock juror gender, expert gender, and testimony complexity in a case involving DNA (deoxyribonucleic acid) evidence. Findings revealed that testimony complexity interacted with expert gender to influence verdict judgments. Participants were unaffected by testimony complexity when the expert was a man, but were more likely to convict when complex testimony was presented by a woman. In support of the heuristic-systematic model, expert gender elicited an effect only in high-complexity conditions—interestingly, this was exclusively the case for male mock jurors. Understanding how jurors cognitively process legal and extra-legal information may help legal actors (e.g. evidence experts, lawyers) communicate evidence and its legal relevance more effectively.

Subjects: Criminology - Law; Criminological Psychology; Applied Social Psychology

Keywords: gender; testimony complexity; DNA evidence; verdict

ABOUT THE AUTHORS
Evelyn M. Maeder is Director of the Institute of Criminology and Criminal Justice at Carleton University, and is cross-appointed with the Department of Psychology. In the Legal Decision-Making Lab, Maeder and her students—including Laura A. McManus, Kendra J. McLaughlin, Susan Yamamoto, and Hannah Stewart—examine a wide array of issues in both the US and Canadian legal systems, in particular discrimination in the law. The goal of this work is to help legal practitioners, scholars, and policy-makers better understand the experiences of marginalized groups by drawing attention to extra-legal factors that can unduly influence decision-making.

PUBLIC INTEREST STATEMENT
Forensic evidence can bring more objectivity to criminal trials, but jurors might be biased by stereotypes regarding forensic science as a “male” profession. In this study, we created a simulated trial to test whether the gender of an expert witness would influence juror judgments depending on whether the testimony was complicated or easy to understand. We found that jurors voted guilty more often when a woman communicated complicated DNA evidence compared to the same testimony from a man, but no such gender-based differences existed for more simplistic testimony. Among male jurors, when the testimony was complex, the female expert was viewed more favorably than her male counterpart in terms of competence, character, and education. Although an expert’s gender is legally irrelevant to the testimony he/she gives, these findings imply that jurors might make different conclusions depending on the expert’s gender and testimony complexity.
1. Introduction
Scientific forensic evidence has transformed legal proceedings within the criminal courtroom. DNA (deoxyribonucleic acid) evidence has been referred to as “the single greatest advance in the search for the truth” and is considered to have the power to both exonerate the innocent and condemn the guilty (People v. Wesley, 1994). DNA is a long molecular chain containing an individual’s unique genetic code, which can be collected from blood, saliva, and other bodily substances (see Miller Albertson, 2009). While DNA evidence is often complex and may not be easily understood by laypersons, its presence in a case can nonetheless influence mock jurors’ decision-making (Lieberman, Carrell, Miethe, & Krauss, 2008). Mock jurors tend to convict more often, and view the victim as more believable, when the Prosecution presents DNA evidence (Golding, Stewart, Yozwiak, Djadali, & Sanchez, 2000).

Despite the persuasive effect of DNA evidence, a host of researchers have demonstrated that extra-legal variables (e.g. witness and juror characteristics) can also nonetheless interfere with juror decision-making (for a comprehensive meta-analytic review, see Devine & Caughlin, 2014). In particular, researchers have implicated gender as a relevant variable in understanding the effectiveness of expert testimony (McKimmie, Newton, Terry, & Schuller, 2004; Schuller, Terry, & McKimmie, 2001, 2005). Such extra-legal variables are especially relevant when the presented evidence and legal information are contradictory or confusing to jurors (Kovera, McAuliff, & Hebert, 1999). Ivković and Hans (2003) have demonstrated that both testimony complexity and characteristics of experts (e.g. personality) affect how credible jurors perceive the experts to be, and consequently their evaluation of the evidence. The purpose of this study was to examine the impact of testimony complexity in criminal cases featuring DNA evidence, and whether expert and juror gender affected decision-making. Understanding how jurors process information may help legal actors (e.g. evidence experts, lawyers) communicate evidence and its legal relevance more effectively. Moreover, insight into how jurors form their verdicts in criminal trials can aid in the development of strategies to reduce juror bias and facilitate a fair trial.

1.1. DNA evidence
Physical forensic evidence is, for the most part, reliable and valid (Lieberman et al., 2008). However, Lieberman et al. (2008, p. 31) cautioned that DNA might serve as the “gold standard” of forensic evidence, but it is by no means the infallible, “platinum” standard. Saks and Koehler (2005) reported that among 86 instances of wrongful conviction—in which DNA evidence later exonerated the defendant—forensic testing errors were responsible in 63% of those cases, and 27% were owing to faulty forensic science testimony. Hence, while DNA evidence has been a powerful source of exonerating the wrongfully accused, virtually any form of physical forensic evidence can be misinterpreted in legal trials due to scientific and human testing errors (Lieberman et al., 2008; Saks & Koehler, 2005). Unfortunately, DNA evidence can be so powerful as to influence juror decisions even when it is not presented. Lieberman et al. (2008) showed that participants sometimes misused blood type evidence (in which the defendant’s blood type matched that found at the crime scene) and erroneously reported that DNA evidence had been presented. Legal practitioners have recently become concerned about the so-called “CSI Effect”, whereby jurors will have unreasonable expectations about sophisticated forensic science, producing fewer convictions (Roane, 2005; Tyler, 2006). While findings on the CSI effect have thus far been mixed (Podlas, 2006; Schweitzer & Saks, 2007; Shelton, Kim, & Barak, 2006), there remains an air of perceived superiority surrounding DNA evidence (Lieberman et al., 2008).

Flawed collection and interpretations of DNA evidence are even more concerning because despite the potential for error, jurors indeed tend to consider it a foolproof indicator of guilt or innocence. For instance, Lieberman et al. (2008) found that participants rated DNA evidence as 95% accurate, 94% persuasive as an indicator of guilt, and overall more accurate than fingerprint or hair fiber evidence. These attitudes extended to verdict decisions, both in terms of acquittal and conviction when evidence was exonerating or incriminating, respectively. Lieberman et al. (2008) also investigated the impact of cross-examination pointing out the limitations of DNA evidence. Although there was no main effect of cross-examination on verdicts, it was successful in alerting jurors to potential errors...
in some instances. Specifically, they manipulated the cross-examination technique, such that either the expert's credibility was challenged, or the evidence itself was questioned regarding potential contamination and other such shortcomings. Additionally, half of the trials featured a lab that was described as being more reliable than the other (e.g. not affiliated vs. directly affiliated with the police department). In trials featuring evidence-focused cross-examination, jurors convicted more often when the lab was described as reliable than when it had questionable reliability. For the unreliable lab, the evidence focused cross-examination decreased perceived accuracy of the DNA evidence. Therefore, it seems that expert testimony has the potential to remedy juror misconceptions of DNA evidence when jurors are provided with accurate and specific information about the process, rather than expert characteristics. However, this does not preclude the possibility that extra-legal variables, such as features of the expert, can interfere with such benefits.

Legal experts and the judicial system tend to assume that jurors are willing and able to systematically process trial information (Manaster, 2013). While jurors may possess the motivation to arrive at a fair verdict, they may not comprehend the content of intricate expert testimonies. Instead, jurors sometimes rely on pre-existing biases, beliefs, and expectations external to the legal content. Researchers have shown that jurors likely use one of two cognitive routes to process information: systematic processing of information and heuristic processing of information (Chaiken, 1980; Levett & Kovera, 2009). Systematic processing involves the methodical and analytical consideration of all information relevant to forming an opinion. Conversely, heuristic processing involves considering a more limited amount of information, using less cognitive effort and using mental “short cuts” to form opinions (Ratneshwar & Chaiken, 1991). The use of one form of cognitive processing vs. the other is largely dependent on the individual's comprehension of the information they are using; people tend to rely on heuristic processing when they do not understand information (Chaiken, 1980). To the aim of ensuring maximal comprehension of evidence quality, it is necessary to understand what circumstances can foster effective expert testimony.

1.2. Expert testimony

In light of recent developments in forensic technologies and research, forensic specialists (e.g. toxicologists, behavioral scientists) have become prevalent expert witnesses in court. An expert witness is any professional who has specialized knowledge, training, and experience in a particular field (Graham, 2011). Expert witnesses are often asked to present, explain, and interpret complex evidence such as DNA to the court with language that can be understood by non-experts (Graham, 2011). These witnesses can be called by either the prosecution or defense, and may present direct testimony, as well as be subject to cross-examination by the opposing attorney. This testimony can be presented at any point during the prosecution or defense’s case. Importantly, expert testimony must meet the Daubert standard, which specifies that “(1) the experts’ specialized knowledge will help the trier of fact [jury] to understand the evidence, (2) the testimony is based on sufficient fact or data, (3) the testimony is based on reliable principles and methods, (4) the expert has reliably applied the principles and methods to the facts of the case” (Daubert v. Merrell Dow Pharmaceuticals Inc., 1993). The criteria include considerations of falsifiability, peer review and publication, known or potential rate of error and existence and maintenance of standards, and general acceptance of the theories intended for presentation (Solomon & Hackett, 1996). The Daubert standard also requires that the trial judge ensure that the evidence and expert testimony meet the aforementioned tenets during juror selection (Solomon & Hackett, 1996). Expert witnesses thus hold the vital role of explaining complex evidence to jurors. However, evidence complexity and the manner of explanation have implications for how jurors process the information and ultimately reach a verdict (Cooper, Bennett, & Sukel, 1996; Lieberman et al., 2008).

1.2.1. Testimony complexity

Studies assessing jurors’ abilities to evaluate evidence have found that jurors' capacity to process complex evidence is often limited (Cecil, Hans, & Wiggins, 1991). Cooper et al. (1996) evaluated the impact of complex expert testimony in the domain of biochemistry on juror decision-making and found that jurors were more likely to rely on cues beyond the content of the evidence when
assessing complex testimony. Specifically, when the expert testimony was easily understood, the researchers found no differences in verdicts as a function of expert credentials (e.g. degree from a prestigious vs. obscure institution). However, when the expert testimony was complex, mock jurors relied on expert credentials as the basis for their judgments. Mock jurors were more likely to side with the plaintiff when the expert with high credentials presented the complex testimony, compared to the expert with modest credentials. Cooper and Neuhaus (2000) found that testimony complexity about polychlorinated biphenyls (PCBs) and its proximal or causational relation to cancer (i.e., high-complexity testimony that contained technical jargon vs. low-complexity testimony that contained a greater proportion of lay terms rather than scientific terms) limited jurors’ ability to centrally process information, leading jurors to be more affected by heuristic cues. When presented with complex testimony, mock jurors were more likely to discount the expert witness with strong credentials and high pay as a result of perceiving him as a “hired gun”, diminishing his effectiveness as a witness. Mock jurors were unaffected by source characteristics in the low complexity conditions. These findings suggest that under cognitively challenging conditions, jurors are likely to shift from systematic to heuristic processing wherein juror biases are likely to have their greatest effect. Hence, research shows that not only do content-related factors (e.g. quality of the evidence; Kovera et al., 1999) influence jurors’ receptivity of expert testimony, but so can source-related factors (e.g. appearance of the witness; Memon & Shuman, 1998).

1.2.2. Expert gender

There is evidence that expert gender is one source characteristic that jurors use to evaluate expert credibility when processing evidence peripherally (Schuller & Cripps, 1998). Research investigating the heuristic role of gender bias in jury trials has found support for gender-domain congruence, in which experts who are men are regarded as more persuasive in male-dominated domains (e.g. construction industry) and experts who are women are perceived as more qualified in female-dominated domains (e.g. clothing industry; Schuller et al., 2001). McKimmie et al. (2004) demonstrated that when an expert testified in a domain consistent case (e.g. automobile business for men, cosmetic business for women), the expert was seen more favorably and therefore had a greater influence on mock jurors’ final judgments. Applied to scientific testimony, these findings suggest that gender role stereotypes may differentially affect the impact of scientific expert testimony on trial outcomes. Researchers have shown that despite female candidates being preferred 2:1 over similarly qualified male candidates, women are significantly underrepresented in science and math focused careers (Williams & Ceci, 2015; Young, Rudman, Buettner, & McLean, 2013). As Ceci, Ginther, Kahn, and Williams (2014) have noted, society appears to hold different expectations about career appropriateness on the basis of gender. Guimond and Roussel (2001) found that men are perceived as better in science and math than women, whereas women are viewed as better in social sciences and language than men. Additionally, research shows that the widespread cultural belief that men are more talented and interested in sciences (i.e., gender-science stereotype) is also a genetically heritable implicit attitude. Specifically, Cai, Luo, Shi, Liu, and Yang (2016) have tested twins’ implicit and explicit gender-science stereotypes and demonstrated that shared genetic (as well as non-shared environmental factors) predict both implicit and explicit gender-science attitudes. Given the scientific nature of DNA evidence, forensic sciences may be more likely to be viewed as stereotypically-male, and therefore gender-domain congruent for male experts.

It seems that both the gender of the expert and the manner in which the expert witness explains the forensic evidence can have significant implications on how well jurors understand the evidence, which might in turn affect their verdict. Drawing on Cooper et al. (1996) findings, Schuller et al. (2005) sought to examine the influence of expert gender and testimony complexity in a civil trial involving a construction-related violation. As expected, gender impacted on mock jurors’ receptivity of the testimony, but only when the evidentiary information was complex. Specifically, when expert testimony was complex, higher damages were awarded to the plaintiff when the expert was male rather than female. These findings are consistent with those of Schuller et al. (2001), who demonstrated expert gender to be a marker of source credibility when there was gender-domain congruence. Interestingly, mock jurors tended to favor the female expert in the low complexity condition,
implying expert gender influenced decision-making differently depending on jurors’ ability to process evidentiary information. Schuller et al. (2005) posit that gender role stereotypes about the experts’ language and presentation style may have impacted mock jurors’ evaluations. The more technical language employed by the woman expert in the complex condition may have been perceived as incongruent with the characteristics typically expected of a witness who is a woman. Consistent with this notion, McKimmie, Newton, Schuller, and Terry (2012) found that mock jurors regarded complex testimony to have a masculine orientation, whereas simple testimony was seen as being more feminine in orientation. Not surprisingly, expert testimony was demonstrated to be most influential when an expert used gender-congruent language, primarily when the expert was female. That is, the female expert was seen as less persuasive when the testimony was presented in a complex, rather than simplified manner. These findings demonstrate that gender stereotypes may extend beyond an expert’s domain of expertise and operate differentially based on jurors’ abilities to evaluate the content of a message.

1.2.3. Juror gender
Gender biases affecting jurors’ perceptions of expert witnesses are not limited to gender role stereotypes. Studies assessing gender differences in decision-making have demonstrated that male and female jurors process evidentiary information differently (Pozzulo, Dempsey, Maeder, & Allen, 2010). Findings from Meyers-Levy and Maheswaran (1991) suggest that while women in the role of mock jurors tend to engage in an effortful and itemized analysis of the evidence, men tend to acquire information heuristically by considering the case as a whole. Assessing the influence of gender on jurors’ evaluations of expert witnesses, Schuller and Cripps (1998) found that clinical expert testimony pertaining to a battered women’s case had more of an impact on men relative to women. For men, exposure to a woman’s expert testimony produced higher believability and lower responsibility ratings for the defendant. In contrast, while women perceived the female expert more favorably, these impressions had little effect on jurors’ decisions. These findings are consistent with the literature on gender-domain congruence and may indicate that jurors perceive women to be better informed than men regarding these kinds of cases (Schuller et al., 2001). These findings also suggest that men may place more weight on expert testimony than women. However, more recent findings suggest that these results may be limited to expert psychological testimony. Research investigating the influence of juror gender in cases with physical evidence has demonstrated no gender differences in juror decision-making (McKimmie et al., 2004; Schuller et al., 2001, 2005).

1.3. Current study
The present study was designed to examine the influence of gender and testimony complexity on juror judgments in a case involving DNA evidence. Specifically, we sought to understand whether expert gender operated as a heuristic cue when testimony regarding DNA evidence was complex, and whether juror gender had any influence in the decision-making process. The study employed a 2 (juror gender: man, woman) × 2 (expert gender: man, woman) × 2 (testimony complexity: low, high) design, creating eight conditions. Drawing from the literature on models of persuasion (e.g. Chaiken, 1980; Ratneshwar & Chaiken, 1991) several hypotheses were developed.

1.3.1. Hypothesis 1: Main effect
Given the gender effects that Schuller et al. (2005) reported, we expected a main effect of expert gender, such that the male expert would be viewed more positively than the female expert as a result of satisfying the gender domain congruence for forensic science (McKimmie et al., 2004). Given the null effects found in the juror gender literature and case domain, we did not anticipate any effects of mock juror gender (Baskin & Sommers, 2010; McKimmie et al., 2004; Schuller et al., 2001, 2005).

1.3.2. Hypothesis 2: Interaction effect
We expected that testimony complexity and expert gender would interact to affect mock jurors’ perceptions of the expert and their verdict decisions. Based on the findings of Cooper et al. (1996), we anticipated that when expert testimony was complex, mock jurors would rely more on heuristic cues (such as gender) in evaluating the expert, such that the male expert would be viewed more
positively and would yield more guilty verdicts. In contrast, we did not expect expert gender to affect mock jurors' evaluations in the low complexity condition, as jurors were expected to be better able to use systematic processing.

2. Method

2.1. Participants
Data were collected from two jury-eligible samples (i.e., Canadian/American citizens, over the age of 18 with no prior indictable/felony convictions) for a total of 170 participants (17 participants who failed one or more manipulation checks were excluded from the sample; see below). Sixty-three participants were undergraduate students recruited from a Canadian university. Student participants (men = 30.2%, women = 69.8%) had a mean age of 20.4 (SD = 4.4). Forty-five percent of the student sample self-identified as Caucasian, 23.8% as Black, 12.7% as Middle Eastern, 7.9% as Asian, 1.6% as Aboriginal-Canadian and 9% identified as another race. One hundred and seven participants were community members recruited online through Amazon's Mechanical Turk. The community sample (men = 43%, women = 57%) had a mean age of 36.3 (SD = 13.6). This sample was predominantly Caucasian (78.3%), with 8.7% self-identifying as Black, 7.5% as Asian, 4.7% as Hispanic and .8% as another race. Participants in the student sample received course credit for their participation, while the community sample received up to $2.00 based on successful completion of attention checks embedded within the study. Six students and 11 community members were dropped from analyses for failing to correctly answer the question “What was the gender of the expert witness in the trial?”.

2.2. Materials and procedure

2.2.1. Trial transcript
Participants’ responses were collected through the online survey tool Qualtrics. Participants had an average completion time of 22 min. Jury-eligible participants read through one of four descriptive accounts of a second-degree murder case, adapted from Lieberman et al. (2008). The case scenario provided a general overview of the criminal incident, such that the victim was stabbed in the park while on a jog. It also featured the testimony/cross-examination of the prosecution’s expert witness (a University of California graduate that had been overseeing the DNA Identification Unit at CellCode for the past 7 years) and defense’s eyewitness (a custodian for the city of Los Angeles working in Colonial Park at the time of the attack). In each scenario, the gender of the expert witness (man, woman) and complexity (low, high) of the testimony varied in the case description.

Testimony complexity was manipulated by using more technical language in the high complexity version (e.g. “Short tandem repeats are short alleles, making them less prone to being affected by degradation, and have slower rates of mutation so that the fragments are more stable in the genome”), and simplified language in the low complexity version (e.g. “Short tandem repeats are very small, so are less likely to be damaged or changed over one’s life”). A pilot study (N = 58) was conducted to ensure there were significant differences between the perceived scientific intricacy (e.g. language complexity) of the expert’s testimony. Significant differences were found between participants’ comprehension of the low complexity testimony (M = 7.00, SD = 1.86) compared to the high complexity testimony (M = 5.23, SD = 2.74; F(1, 56) = 9.05, p < .001). In the final case scenario, the gender of the expert was manipulated using names (e.g. Frank Miller for the male expert, Francine Miller for the female expert), titles, and pronouns, as well as photographs that were pilot-tested (N = 35) to match on perceived age (t(34) = 0.00, p = 1.00) and attractiveness (t(34) = .550, p = .82). In the final case scenario, one photo of the prosecution’s expert witness (man or woman) and one photo of the defense’s eyewitness (man) accompanied the testimony/cross-examinations. To help inform mock juror judgments, participants were also provided with the California Penal Code criteria for the charge, as well as instructions regarding the burden of proof and reasonable doubt.
2.2.2. Juror questionnaire
Participants then completed a number of opinion related measures in the form of a juror questionnaire. The questionnaire included a dichotomous judgment of guilt (guilty/not guilty), as well as the extent to which the DNA evidence influenced their decision ranging from 1 (not at all) to 9 (very much). Participants indicated how compelling and credible they perceived the DNA evidence to be on a scale from 1 (not at all) to 9 (very much).

To assess participants’ impression of the prosecution’s expert witness, participants rated the expert on competence (i.e., believability, honesty, trustworthiness, credibility, clarity, comprehensibility, helpfulness, professionalism, and influence) and character (i.e., likeability, attractiveness, persuasiveness) on 9-point rating scales from 1 (not at all) to 9 (very much). Subjective questions regarding the expert’s level of education and experience conducting DNA analyses were assessed separately on scales from 1 (minimally educated/experienced) to 9 (very educated/experienced).

Given the various domains within these measures, we assessed the dimensionality of the items so as to combine questions that substantially overlapped. We conducted a principal components analysis with a direct oblimin rotation, first using the default Kaiser criterion of eigenvalues greater than one (i.e., revealing components that account for more variance than any single item). This criterion resulted in a three-component structure. However, components were only retained if at least three items showed loadings of .32 or greater, as this value represents roughly 10% of overlap in variance (Tabachnick & Fidell, 2001). Examining the component loadings demonstrated that only two items—those pertaining to expert experience and expert education—loaded onto the second component, and so they were dropped. Therefore, results were supportive of a two-component structure, which accounted for a cumulative 66.5% of the variance. The first component contained items pertaining to expert competence (i.e., believability, honesty, trustworthiness, credibility, clarity, comprehensibility, helpfulness, professionalism, and influence). The second component was comprised of items about the expert’s character (i.e., likeability, attractiveness, persuasiveness).

The expert competence component, which contains 9 items, demonstrated strong internal consistency ($\alpha = .95$), with strong item-total correlations ($> .60$), and Cronbach’s $\alpha$ would decrease with deletion of any item. Similarly, the expert character component, which contained three items, showed strong internal consistency, $\alpha = .82$, strong item total correlations ($> .60$), and a decrease in Cronbach’s $\alpha$ with any item deletion. Based on these results, we decided to analyze the components of expert competence and expert character separately. In order not to lose the information from the education and experience measures, we analyzed these items as individual variables.

3. Results

3.1. Verdict
Overall, our results revealed a fairly even verdict split, with 47% ($n = 80$) voting not guilty, and 53% ($n = 90$) voting guilty. To determine whether students and community members would differ in their verdict decisions, we conducted a Chi-Square analysis. This test revealed no significant relationship between sample type and verdict, $\chi^2 (1, N = 170) = 1.35, p = .25, \nu = .09$. As such, we collapsed across sample type for the remaining verdict analyses.

In order to determine whether juror gender, expert witness gender, and testimony complexity would influence verdict decisions alone or in combination, we conducted a 2 (juror gender: man, woman) × 2 (expert witness gender: man, woman) × 2 (testimony complexity: low, high) × 2 (verdict: guilty, not guilty) hierarchical log-linear analysis. See Table 1 for a breakdown of participant gender in each condition.

This analysis revealed a non-significant expert gender x testimony complexity x verdict interaction, $\chi^2 (1, N = 170) = 2.96, p = .08$. However, follow-up analyses revealed that when testimony complexity was high, a significant effect of gender emerged, $\chi^2 (1, N = 80) = 5.21, p = .02, \nu = .26$, such
that participants were significantly more likely to vote guilty when the expert was a woman (69%) as compared to when he was a man (44%). In contrast, when testimony complexity was low, there was no effect of expert gender, $\chi^2(1, N = 90) = .05, p = 1.00, \eta^2 = .02$. See Table 2 for a summary of these results.

### 3.2. Perceptions of the DNA evidence

In general, participants rated the DNA evidence very highly—they found the evidence very compelling ($M = 7.21, SD = 1.61$), indicated that it strongly influenced their verdict decisions ($M = 7.13, SD = 1.71$), and believed it to be very credible ($M = 7.21, SD = 1.76$). These perceptions did not vary as a function of sample type, ($t$ values between −.02 and −1.63, all $p$’s > .10), and so further analyses were collapsed across this variable.

A correlation analysis revealed that our DNA evidence ratings (i.e., the degree to which participants felt that the evidence was compelling, influenced their verdict, and was credible) were associated ($r$’s ranging from .71 to .81; all $p$’s < .001), and so we conducted a multivariate analysis of variance (MANOVA) using our DNA evidence rating variables as the dependent measures, and juror gender (man, woman), expert gender (man, woman), and testimony complexity (low, high) as the independent variables. This analysis revealed no significant multivariate effects, and only one significant univariate effect—a main effect for expert witness gender on the degree to which participants felt that the DNA evidence influenced their verdicts, $F(1, 168) = 4.73, p = .03, \eta^2 = .03$. Participants who saw testimony from a female expert witness ($M = 7.41, SD = 1.53$) indicated that the DNA evidence influenced their verdicts more than those who saw testimony from a male expert witness ($M = 6.88, SD = 1.82$).

### 3.3. Perceptions of the expert witness

In general, participants rated the expert witness quite highly (means ranged from 6.46 to 8.01). None of these perceptions varied as a function of sample type ($t$ values ranging from .24 to 1.37, all $p$’s > .10), and so we collapsed across sample type for further analyses of these variables.

#### 3.3.1. Expert competence

We observed a significant juror gender x expert gender interaction on this subscale, $F(1, 168) = 7.86, p = .006, \eta^2 = .05$. Follow-up analyses revealed that for male jurors, there was a significant effect of expert witness gender, $t(63) = −2.94, p = .005$, such that the male expert ($M = 6.69, SD = 1.55$) was
seen as less competent than the female expert ($M = 7.70, SD = 1.15$). There was no effect of expert witness gender on female jurors, $t(101) = .92, p = .36$. See Figure 1 for a depiction of these results.

3.3.2. Expert character

For this component of expert perceptions, we observed a significant main effect of expert witness gender, $F(1, 168) = 4.39, p = .04, \eta^2 = .03$. This was qualified by a significant three-way interaction among expert witness gender, juror gender, and testimony complexity, $F(1, 168) = 4.83, p = .03, \eta^2 = .03$.

Follow-up tests demonstrated that when testimony complexity was low, there was a significant effect of juror gender on expert character perceptions, $F(1, 89) = 5.61, p = .02, \eta^2 = .06$. Male participants ($M = 5.55, SD = 1.56$) provided lower ratings of the expert’s character than did female participants ($M = 6.30, SD = 1.34$).

When testimony complexity was high, a different pattern emerged—specifically, we observed a significant interaction between juror gender and expert gender, $F(1, 79) = 5.32, p = .02, \eta^2 = .07$. Follow-up tests revealed that female jurors did not differ in character assessments of male ($M = 6.14, SD = 1.26$) and female ($M = 5.74, SD = 1.89$) expert witnesses, $t(49) = .91, p = .37$, in conditions of high complexity. However, male jurors provided higher ratings of character to the female ($M = 6.56, SD = 1.23$) expert as compared to her male counterpart ($M = 5.14, SD = 6.56$), $t(26) = −2.03, p = .05$.

3.3.3. Experience

For this measure, a three-way interaction among expert gender, juror gender, and testimony complexity was observed, $F(1, 167) = 8.05, p = .005, \eta^2 = .05$. Follow-up tests showed that when testimony complexity was low, there were no significant effects of juror gender or expert gender. However, when testimony complexity was high, there was a significant expert gender x juror gender interaction, $F(1, 79) = 4.40, p = .04, \eta^2 = .06$. Simple effects tests demonstrated that when a male expert testified with high complexity, there was no effect of juror gender on perceptions of his experience, $t(39) = −.71, p = .48$. However, when a female expert testified with high complexity, juror gender was significantly related to perceptions of her experience, $t(36) = 2.03, p = .05$, such that men ($M = 8.53, SD = .64$) rated her as more experienced than did women ($M = 7.43, SD = 2.02$).
3.3.4. Education
We observed a significant juror gender x testimony complexity interaction on participants’ perceptions of the expert’s education, \( F(1, 168) = 3.87, p = .05, \eta^2 = .02 \). This was qualified by a three-way interaction among expert gender, juror gender, and testimony complexity, \( F(1, 168) = 8.50, p = .004, \eta^2 = .05 \). Follow-up tests showed that when testimony complexity was low, no significant effects of expert witness or juror gender were observed. However, when testimony complexity was high, we observed a significant expert gender x juror gender interaction, \( F(1, 79) = 6.28, p = .01, \eta^2 = .08 \). Simple effects tests demonstrated that for male jurors, there was a significant main effect of expert gender, \( t(26) = −2.12, p = .04 \), such that they perceived the female expert (\( M = 8.27, SD = .88 \)) as more educated than the male expert (\( M = 7.46, SD = 1.13 \)). There was no significant effect of expert witness gender on female jurors.

4. Discussion
The purpose of this study was to examine the effects of testimony complexity on juror decision-making in a case involving DNA evidence, and whether expert or juror gender influenced use of that testimony. Participants read a description of a second-degree murder case, adapted from Lieberman et al. (2008), which featured testimony/cross-examination of the prosecution’s expert witness. In line with previous research demonstrating the persuasiveness of DNA evidence, participants in this study generally rated it as credible, compelling, and influential in their verdict decisions. Notably, those who read testimony from a female expert rated the evidence as more influential in their eventual verdict. This finding deviates from earlier work on gender-domain congruency (McKimmie et al., 2004; Schuller et al., 2001), which posits that male experts would be more effective in this traditionally-male field (science). Given that our experts were pilot-tested to match on age and attractiveness, it appears that mock jurors in this study were simply more persuaded by a female expert. This suggests that forensic science/DNA analysis may not be perceived as a stereotypically-male domain, or that gender-domain congruency was not supported in this particular study.

We compared the student and community samples on each of our dependent measures, and found there to be no significant differences between them. As such, the use of a student sample in this domain may be an appropriate proxy.

We expected a testimony complexity by expert gender interaction effect on perceptions of the expert as well as on verdict decisions. In line with findings from Cooper et al. (1996), we hypothesized that when expert testimony was complex, mock jurors would rely more on heuristic cues (such as gender) in evaluating the expert. In contrast, we did not expect expert gender to affect mock jurors’ evaluations in the low complexity condition. Following findings that Schuller et al. (2005) reported, we also predicted a main effect of expert gender, such that the male expert would be viewed more positively than the female expert (McKimmie et al., 2004). The data did not fully support these hypotheses. We found that when testimony complexity was low, there was no effect of expert gender, but when testimony complexity was high, participants were significantly more likely to vote guilty when the expert was a woman as compared to when he was a man. This runs contrary to earlier work in the area, which demonstrated that male experts were more effective when testimony complexity was high, and female experts were more effective when testimony complexity was low (Schuller et al., 2005). Given the paucity of work in this area, it is difficult to say whether these differences reflect different domains of testimony, the contrast between civil and criminal trials, or some other phenomenon. Future work will be needed in order to properly tease apart the complex relationship between gender and language in this context.

Results, however, did suggest that complex testimony might invite gender biases. We examined the impact of gender and complexity on perceptions of the expert, in terms of perceived competence, character, experience, and education. In conditions with high complexity, we observed a number of interesting gender effects—in particular, male jurors rated the female expert witness more favorably than her male counterpart with regard to competence, character, and education, whereas female jurors did not exhibit expert witness gender effects. With regard to perceived
experience, under conditions of high complexity the male expert was perceived largely the same by both genders, whereas male jurors rated the female expert more favorably than did female jurors. These findings demonstrate a number of heuristic influences under conditions of high complexity, as suggested by earlier research on the heuristic-systematic processing model (Chaiken, 1980).

The male jurors’ preference for the female expert in terms of character, education, experience, and competence deserves further attention. These results are contrary to our expectations, and to previous literature linking science to male stereotypes (Nosek et al., 2009) and gender-domain congruency (McKimmie et al., 2004; Schuller et al., 2001). They may reflect social desirability concerns of our male participants, or they may reflect the shifting of stereotypes based on a growing number of women in science or media portrayals of such (Cavender & Deutsch, 2007; Deutsch & Cavender, 2008). For example, Cavender and Deutsch (2007, p. 72) analyzed the content of several episodes of Crime Scene Investigation (CSI) and reported that “CSI’s female characters have essentially the same duties and abilities as their male counterparts,” but also that the female forensic investigators may have “special women’s insights”. These depictions of competent (although still stereotypically female) women in the forensic science context may be improving perceptions of female experts in this area. Other research has shown that perceptions of various evidence types are influenced by crime television, although verdicts are not directly affected (e.g. Maeder & Corbett, 2015; Schweitzer & Saks, 2007). However, it must be noted that content analyses of crime-television shows demonstrate that they are far from reflective of reality (Ley, Jankowski, & Brewer, 2012). Aside from television, it may be the case that our participants (many of whom were university students) are accustomed to seeing or even working with women scientists and thus do not hold explicit (i.e., deliberate) gender-stereotypic views of science and women (see Miller, Eagly, & Linn, 2014). Miller et al. (2014) suggest that gender-science stereotypes are culturally variable and therefore have potential to change. Additional research should investigate whether exposure to female forensic scientists through personal relationships or media portrayals is related to perceptions of expert testimony, and whether women in other scientific domains (e.g. chemistry or physics) are regarded as favorably. Given the notable number of female forensic experts on television, perhaps there is less gender incongruence than predicted from the base rates of women in the science professions. For now, these findings shed light on a possible shift in men’s perceptions of women in science, at least in the courtroom.

4.1. Limitations

This study featured some notable limitations. First, all juror simulation research has some limitations to ecological validity. In particular, this study used a written trial scenario rather than a video or live trial, and it also did not feature a deliberation component. However, researchers have examined the influence of trial medium on juror decision-making, and have shown that written descriptions yield comparable results compared with more realistic presentations (Bornstein, 1999; Pezdek, Avila-Mora, & Sperry, 2010). Further, individual juror judgments are widely recognized as an important deciding factor in final verdict outcome (Kalven & Zeisel, 1966). Additionally, we did examine juror decision-making in both community and student samples—for which we did not observe any differences—bolstering the generalizability of our findings. Sears (1986) has argued that students’ decision-making can differ from that of community members, and Bray and Kerr (1982) have cautioned that students do not often serve as jurors.

Perhaps more significantly, some of our findings fell short of traditional cutoff values, and so results should be interpreted with caution. Also of note is that participants may have been wary of expressing stereotype congruent or sexist views about the female expert witness. However, most of these effects occurred in the low complexity conditions, for which previous studies suggest that gender differences might not be expected, due to the stronger availability of systematic processing.
Another notable limitation is the imbalance of male and female participants in each condition (see Table 1). Although this imbalance existed in both high- and low-complexity conditions, it could partially explain why we were unable to uncover 2- or 3-way interactions with participant gender.

Finally, these data came from an online sample, which limited our control over participants' study environments. The use of manipulation checks helps to filter out participants who did not attend closely to the trial stimulus. Further, several studies support the conclusion that Mechanical Turk tends to yield reliable data (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Gardner, Brown, & Boice, 2012; Paolacci, Chandler, & Ipeirotis, 2010).

5. Conclusion

In general, this study demonstrated a number of interesting gender effects in the context of a murder trial involving the presentation of DNA evidence. As posited by the heuristic-systematic model, these effects were found largely in conditions involving highly complex testimony, suggesting that gender may play a significant role in the courtroom when jurors are less able to comprehend the information that is presented to them. This has a number of implications for the delivery of expert testimony in this domain, and for jury selection as well, since (unexpectedly) male jurors rated the female expert more favorably than her male counterpart. Future research should continue to investigate perceptions of experts as a function of gender and testimony complexity, in order to determine whether these effects are applicable beyond the context of DNA evidence. Given the prominent role that DNA testimony now plays in our court system, these findings shed light on the potential roles of gender and language in criminal trials.

References

Baskin, D. R., & Sommers, I. B. (2010). Crime-show-viewing habits and public attitudes toward forensic evidence: The “CSI Effect” revisited. Justice System Journal, 31, 97–113. doi:10.1080/0098261X.2010.7067956

Borsboom, B. H. (1990). The ecological validity of jury simulations: Is the jury still out? Law and Human Behavior, 23, 75–91. doi:10.1023/A:1022326807441

Bray, R. M., & Kerr, N. L. (1982). Methodological considerations in the study of the psychology of the courtroom. In N. L. Kerr & R. M. Bray (Eds.), The psychology of the courtroom (pp. 287–323). San Diego, CA: Academic Press Inc.

Cai, H., Luo, Y. L., Shi, Y., Liu, Y., & Yang, Z. (2016). Male = science, female = humanities: Both implicit and explicit gender-science stereotypes are heritable. Social Psychological and Personality Science, 7, 412–419. doi:10.1177/1948550615627367

Cavender, G., & Deutsch, S. K. (2007). CSI and moral authority: The police and science. Crime, Media, Culture, 3, 67–81. doi:10.1007/s12149-007-9200-x

Ceci, J. S., & Sloop, V. (1991). Citizen comprehension of difficult issues: Lessons from civil jury trials. The American University Law Review, 40, 727–774. Retrieved from http://scholarship.law.cornell.edu/fulPERATURE

Ceci, S. J., & Wunsch, R., & Williams, W. M. (2014). Women in academic science: A changing landscape. Psychological Science in the Public Interest, 15, 75–141.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. Journal of Personality and Social Psychology, 39, 752–766. doi:10.1037/0022-3514.39.5.752

Cokely, E. T., Galesic, M., Schultz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. Judgment and Decision Making, 7, 25–47.

Cooper, J., & Neuhaus, I. M. (2000). The “Hired Gun” effect: Assessing the effect of pay, frequency of testifying, and credentials on the perception of expert testimony. Law and Human Behavior, 24, 149–171. doi:10.1023/A:1005476618439

Cooper, J., Bennett, E. A., & Suek, H. L. (1996). Complex scientific testimony: How do jurors make decisions? Law and Human Behavior, 20, 379–394. doi:10.1023/A:1008050758724

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

Douglas v. Merrell Dow Pharmaceuticals. (1993). 509 U.S. 579.

Deutsch, S. K., & Cavender, G. (2008). CSI and forensic realism. Journal of Criminal Justice and Popular Culture, 15, 34–53.

Edgell, M. L., & Caughlin, D. E. (1992). Mechanical Turk tends to yield reliable data (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Gardner, Brown, & Boice, 2012; Paolacci, Chandler, & Ipeirotis, 2010).
Tyler, T. R. (2006). Viewing CSI and the threshold of guilt: Managing truth and justice in reality and fiction. The Yale Law Journal, 115, 1050-1085. http://dx.doi.org/10.2307/2045564

Williams, W. M., & Ceci, S. J. (2015). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. Proceedings of the National Academy of Sciences, 112, 5360-5365. http://dx.doi.org/10.1073/pnas.1418878112

Young, D. M., Rudman, L. A., Buettner, H. M., & McLean, M. C. (2013). The influence of female role models on women’s implicit science cognitions. Psychology of Women Quarterly, 37, 283-292. doi:10.1177/0361684313482109