Differences in Expert Witness Knowledge: Do Mock Jurors Notice and Does It Matter?

Caroline T. Parrott, PhD, Tess M. S. Neal, PhD, Jennifer K. Wilson, MA, and Stanley L. Brodsky, PhD

The knowledge of experts presumably affects their credibility and the degree to which the trier of fact agrees with them. However, specific effects of demonstrated knowledge are largely unknown. In this experiment, we manipulated a forensic expert's level of knowledge in a mock-trial paradigm. We tested the influence of low versus high expert knowledge on mock juror perceptions of expert credibility, on agreement with the expert, and on sentencing. We also tested expert gender as a potential moderator. Knowledge effects were statistically significant; however, these differences carried little practical utility in predicting mock jurors' ultimate decisions. Contrary to the hypotheses that high knowledge would yield increased credibility and agreement, knowledge manipulations influenced only perceived expert likeability. The low-knowledge expert was perceived as more likeable than the high-knowledge counterpart, a paradoxical finding. No significant differences across expert gender were found. Implications for conceptualizing expert witness knowledge and credibility and their potential effects on juror decision-making are discussed.


Knowledge and competence are characteristics that serve a key role in human interactions. Interpersonal effectiveness and positive impression management are affected by perceptions of intellectual ability, knowledge, and skill. Knowledge may be communicated through self-proclamation, assertiveness, substantive content, or experience. An expert's acquired knowledge and competence may serve an important role in a courtroom, where lives and livelihoods may hang in the balance. After all, knowledge relates initially to whether a particular professional is retained to testify. The rules of evidence explicitly identify knowledge, experience, training, education, or skill as the practical foundations on which the witness is deemed an expert and permitted to testify by the court.

Unlike other credibility influences, such as confidence or trustworthiness, expert knowledge is mandated by the court in the rules of evidence governing acceptance of a witness as an expert. Thus, an expert witness's knowledge is doubly subjected to scrutiny by the court and the trier of fact. As suggested by previous research, the court's sanction of a witness as expert likely serves as a heuristic to triers of fact in their evaluations of experts' credibility in many cases. Conversely, a witness's expert status has the potential to backfire and create distrust in the form of skepticism that the expert is a hired gun or feelings of comparative inferiority on the part of the trier of fact. The effects of expert qualifications and displays of knowledge during testimony are largely unknown.

Expert Witness Credibility in the Courtroom

An expert witness's credibility has the potential to influence jurors' consideration of their testimony. Both expert knowledge and credibility have been shown to influence disputing parties and third-party decision-makers. Expert witness credibility has been instrumental in verdicts and sentencing recommen-
Constraining Expert Witness Credibility

The Witness Credibility Model (WCM) is a framework that conceptualizes witness credibility as a composite of four factors: confidence, likeability, knowledge, and trustworthiness. The four-factor model is effectively captured by the Witness Credibility Scale (WCS). Numerous studies have used the WCS and validated its usefulness in evaluating perceptions of expert witnesses. Across these investigations, the WCS has held its conceptual strength and demonstrated adequate internal consistency and reliability.

Knowledge in the Courtroom Setting

Expert Qualification

Because of the potential influence of expert witness testimony, experts must be qualified by the courts. In a survey of judges, jurors, lawyers, and experts in civil trials, Champagne et al. found knowledge and expertise to be the most desired characteristics in an expert witness. Moreover, perceptions of knowledge were closely linked to impressive educational credentials and reputation as a leading expert in the field. Legal requirements generally define expertise as acquired through relevant experience, training, knowledge, education, or skill. Expert knowledge may be demonstrated through academic degrees obtained, positions held, particular populations evaluated or treated, professional certifications or licensure, board certification, membership in professional organizations, professional publications, prior court experience as an expert, and honors and awards. Thus, the qualification process becomes almost synonymous with credentialing.

Commons et al. suggest that the way expert witness qualifications are presented to jurors could affect how jurors view the expert and by extension, perhaps how the jurors evaluate the testimony. Hurwitz et al. conducted a language and content analysis of actual trial transcripts. They concluded that the jurors perceived expert witnesses as more credible if the experts presented content related to their credentials or experience (i.e., expertise) and objectivity (i.e., trustworthiness) during expert qualification.

Knowledge on the Stand

In the credentialing procedures for an expert witness, the court treats each of the five characteristics outlined in the Federal Rules of Evidence—experience, education, training, skill, and knowledge—as independently representative of expertise. However, research has shown that experience does not necessarily equate with improved accuracy or knowledge. Thus, to be viewed as expert by the trier of fact (and not just by the rules of evidence), expert witnesses should demonstrate mastery of their craft, conveying their knowledge through testimony. As Champagne and colleagues reported, jurors especially appreciate experts who can make testimony understandable to the lay person and communicate technical information simply and clearly by avoiding or explaining any jargon. Scholars have described how triers of fact may benefit from knowledge woven into a comprehensive story of the evidence. Testimony should accordingly be consistent with commonsense understanding of physical evidence and the testimony of other witnesses. Researchers have also explored jurors’ sensitivity to differences in the quality and presentation style of data cited by expert witnesses, as well as the presence or absence of an expert. However, to our knowledge, no study has experimentally isolated and manipulated level or degree of expert knowledge on the stand to test its influence on decision-making.

Two of the four Witness Credibility Scale factors, likeability and confidence, have been experimentally manipulated and studied in relation to the Witness Credibility Model (Table 1). The main effect of knowledge on expert credibility has yet to receive similar empirical attention. Neal et al. studied expert witness knowledge, but only as it interacted with likeability. That is, they did not isolate knowledge in that study; rather, they varied knowledge and likeability at the same time and studied their interactions rather than their main effects.

They found that likeability and knowledge did interact in the expert witness role, with higher levels of likeability and knowledge being associated with higher credibility. However, they were not able to discern to what degree expert knowledge alone affects perceptions of credibility.

Gender as a Moderator of Perceived Knowledge

Prior research has found inconsistencies in whether the expert’s gender moderates perceptions of expert witness credibility. For example, studies...
have found differential effects based on the experts’ gender, either in favor of men or in favor of women. Other studies have uncovered complex interactions in the ways in which male and female experts are perceived. For example, Neal et al. found that experts who met threshold expectations of likeability and knowledge were not perceived differently based on their gender; however, when they were not likeable or particularly knowledgeable, male experts were perceived significantly more positively and were more persuasive than female experts. We included expert gender as an independent variable in the present study to explore further the relation between expert gender and credibility.

Table 1

<table>
<thead>
<tr>
<th>WCM Factor</th>
<th>Definition</th>
<th>Operational Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likeability</td>
<td>The degree to which an expert is friendly, respectful, kind, well-mannered, and pleasant.</td>
<td>High likeability: consistent use of “we” or “us” when discussing members of the scientific community or humanity as a whole, moderate levels of smiling, modest statements and conclusions (e.g., “relatively certain” or “we do not know everything there is to know in psychology”), consistent eye contact with lawyer and jury, and informal speech (i.e., limited technical jargon and use of surnames of parties in the courtroom). Low likeability: no use of “we” or “us”, no smiling, excessive statements of certainty of conclusions, inconsistent eye contact, use of highly technical jargon, and frequent formal references (e.g., “the client”, “the defendant”).</td>
</tr>
<tr>
<td>Knowledge</td>
<td>The degree to which an expert is perceived to be well informed, competent, or perceptive and to possess or exhibit intelligence, insight, understanding, or expertise.</td>
<td>High knowledge: strong educational credentials (e.g., board certification, history of academic publication in case-relevant area of expertise (educated at Yale, American Board of Forensic Psychology certified, history of relevant publications), solid relevant clinical and research experience (researches risk assessment, has conducted over 100 clinical risk assessments over 14 years), consistent clarity and substantive content of communication, moderate assertiveness (e.g., “as far as I know I’ve never been wrong” when queried about awareness of clinician error), self-proclaimed expertise (e.g., “In my expert opinion...”), and demonstrates familiarity with the case (e.g., multiple interviews with the defendant). Low knowledge: no mention of educational credentials, minimal relevant experience (e.g., little experience or nonrelevant experience [two years as a psychotherapist and no previous risk assessment experience], inconsistent clarity and substantive content of communication, low assertiveness (e.g., “no” when queried about awareness of clinician error), no self-proclaimed expertise, inadequate familiarity with the case (e.g., one short interview with the defendant the week the case went to trial).</td>
</tr>
<tr>
<td>Confidence</td>
<td>The degree of demonstrable self-assurance expert witnesses have in their general ability on the stand.</td>
<td>Low confidence: quivering tone of voice, dysfluencies in speech, vacillating pace of speech, corrections, breaks in the flow of words, postural awkwardness, fixed eye contact, saying “you know” to seek assurance, asking for repetition of questions, and signs of anxiety and nervousness. Medium confidence: moderate and stable tone of voice, clarity in speech, moderately paced speech, willingness to acknowledge a degree of certainty (“I am reasonably certain”), smooth narrative statements, good posture and straight back, comfort and poise, consistent eye contact, accurate hearing, and appropriate responses. High confidence: loud and strong tone of voice, assertive speech and mannerisms, rapidly paced speech, always and all statements (“I am certain”), good posture/leaning forward, high fluency of speech.</td>
</tr>
<tr>
<td>Trustworthiness</td>
<td>Not yet defined as part of the WCM.</td>
<td>Has not yet been operationally defined within the WCM.</td>
</tr>
</tbody>
</table>
The Current Study

Perceptions of expert witness credibility may vary as a function of knowledge presentation. For example, some experts may not deem it necessary to discuss their specialized knowledge once qualified. Others may present displays of their knowledge judiciously and throughout their testimony in an effort to emphasize their expertise. Studying how various demonstrations of expert knowledge influence juror decision-making is a step toward understanding the effectiveness of expert testimony.

In the current study, we focused specifically on the main effect of expert witness knowledge. We sought to examine juror perceptions of expert credibility and varying degrees of expert knowledge. We manipulated expert knowledge as the independent variable (high versus low knowledge) while holding other WCS constructs constant. We expected to find a difference between high- and low-knowledge manipulations of the expert on the following dependent variables: the three other components of credibility—trustworthiness, likeability, and confidence; and sentencing recommendations, as well as agreement with the expert’s opinion on likelihood of future violence. We specifically hypothesized that the very knowledgeable expert (compared with the less knowledgeable counterpart) would be rated significantly higher on credibility outcomes and yield more mock juror agreement with the expert regarding likelihood of defendant future violence and sentencing recommendations. Drawing on the inconsistent findings regarding the effects of expert witness gender on perceptions of credibility in prior research, we explored gender effects in the current study. That is, given the potential interaction of expert gender and knowledge on credibility, we included expert gender as a second independent variable.

Methods

Study Design and Operational Definitions

This study was a 2 (high versus low knowledge) × 2 (male versus female expert witness), between-subjects, factorial design. Thus, the independent variables were knowledge (high versus low) and gender (male versus female). We defined expert knowledge as, “the degree to which an expert is perceived to be well-informed, competent, or perceptive and to possess or exhibit intelligence, insight, understanding, or expertise” (Ref. 10, p 490). A literature review identified components associated with high knowledge, as displayed in Table 2. This conceptualization has been supported in previous work in which the interactions between knowledge and likeability were examined.10

Our operational definition of knowledge included substantive content and clarity of testimony, credentials, relevant experience, self-proclaimed expertise, assertiveness, and familiarity with the case. The specific manipulated conceptions of high and low knowledge, again drawing on Neal et al.,10 are detailed in Table 1.

Participants

Undergraduate psychology students (n = 155) at a large public university participated for course credit. The U.S. Supreme Court decided in Witherspoon v. Illinois37 that jurors who sit on capital murder trials must be death qualified—that is, willing and able to consider capital punishment as a sentencing option. Because our stimulus material was based

<table>
<thead>
<tr>
<th>Table 2 Conceptual Components of Knowledge in Previous research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Substantive content of communication</td>
</tr>
<tr>
<td>Assertiveness</td>
</tr>
<tr>
<td>Clarity of communication</td>
</tr>
<tr>
<td>Educational credentials</td>
</tr>
<tr>
<td>Familiarity with the facts of the case</td>
</tr>
<tr>
<td>Relating testimony content to physical evidence and other witnesses</td>
</tr>
<tr>
<td>Sufficient experience relevant to the content of the communication</td>
</tr>
<tr>
<td>Testimony’s consistency with common sense</td>
</tr>
<tr>
<td>Academic degrees obtained, positions held, populations evaluated or treated, professional certifications or licensure, board certification, membership in professional organizations, professional publications, prior court experience as an expert, and honors and awards</td>
</tr>
<tr>
<td>Self-proclaimed credentials</td>
</tr>
</tbody>
</table>
on the sentencing phase of a capital murder trial, those mock jurors who indicated absolute opposition to the death penalty were excluded from our analyses ($n = 13$), and six mock jurors were removed due to missing data, reducing the total sample size from 155 to 136. Mock jurors who were ineligible due to the Witherspoon criteria were distributed equally across the study conditions and reflected the overall demographic makeup of eligible participants. The gender composition of the sample was 81 percent female and ranged in age from 18 to 43 years (mean ($M$) = 18.76; standard deviation ($SD$) = 2.54). The sample was 79 percent Caucasian, 13 percent African-American, and 8 percent other racial or ethnic backgrounds.

**Stimuli**

We developed four separate videos, each approximately five minutes in length, to match the experimental conditions: male expert witness–high knowledge; male expert witness–low knowledge; female expert witness–high knowledge; female expert witness–low knowledge. Real expert witnesses testified in this mock scenario (rather than actors). When the video opened, the judge described to the mock jurors that the hearing represented the capital sentencing phase of Mr. Jones, a defendant who had already been found guilty of first-degree murder. The judge explained that the only task before the jurors was to decide whether Mr. Jones should be sentenced to death or life in prison. The judge explained the standard for burden of proof before the expert testified. In each video, the expert witness testified under both direct and cross-examination about evaluation of Mr. Jones’ likelihood of future violence. In all conditions, the expert testified to the substantial likelihood that the defendant would reoffend. The video script was adapted from a jury sentencing proceeding used in previous studies. The script was modified to reflect the knowledge manipulations described above and included either a male or female expert matched for age, race, and clothing.

**Procedure and Materials**

Before the study began, Institutional Review Board approval was obtained from the University of Alabama Office of Research Compliance for research with human subjects. Information about the study procedures and details regarding informed consent were provided to participants, and then they viewed a randomly assigned video condition. After watching the video, all participants individually completed the following questionnaires.

**Witness Credibility Scale**

The Witness Credibility Scale (WCS) was used to assess the credibility of the expert. The scale contains 20 bipolar adjectives on a 1- to 10-point Likert scale (e.g., unkind (1) to kind (10); dishonest to honest; shaken to poised). Higher scores indicate higher credibility ratings. The WCS generates an overall credibility rating ($\alpha = 0.96$ in this study) with higher scores indicating higher credibility. The WCS also yields a multidimensional measure of expert credibility defined by four subordinate domains: trustworthiness ($\alpha = 0.95$), confidence ($\alpha = 0.92$), likeability ($\alpha = 0.90$), and knowledge ($\alpha = 0.93$). Given the present study’s interest in how expert knowledge may relate to operational, potentially changeable facets of credibility (e.g., likeability and confidence), expert credibility was assessed at the facet level (i.e., trustworthiness, confidence, and likeability) instead of at the global level (overall credibility) in this study.

**Future Violence Likelihood Rating**

Participants were asked to rate from 1 to 100 percent the likelihood that the defendant would commit future acts of violence. These ratings reflected how believable the participant found the expert who opined that the defendant was likely to commit future violent acts. Thus, this outcome reflects mock jurors’ evaluations of the defendant’s likelihood of engaging in future violence and mock jurors’ agreement with the expert.

**Sentencing Rating**

On Likert-type scales ranging from extremely likely to extremely unlikely, participants rated their likelihood of sentencing the defendant to each of the two available sentencing options: life in prison without parole (LWOP) or death. To create a single continuous sentencing variable, these two Likert-type ratings were converted to standardized $z$ scores. Then, the death penalty $z$ scores were multiplied by $-1$, and the LWOP $z$ scores were multiplied by $+1$. Finally, the two sets of $z$ scores were summed to create a single continuous sentencing variable that conveys both direction and strength. That is, the more negative the score, the more likely the participant would be to assign the death penalty (representing agreement with the expert). The more positive
the score, the more likely the participant would be to assign LWOP (disagreement with the expert).

**Demographics**

A demographics questionnaire elicited participants’ age, gender, and degree of death penalty support.

**Manipulation Checks**

The Knowledge subscale of the Witness Credibility Scale was used as a manipulation check. This subscale comprises five items, including queries about whether the expert seems logical, informed, wise, educated, and scientific (again, \( \alpha = 0.93 \) in this sample).\(^7\) In addition, we included one question about the attractiveness of the expert witness.

Target attractiveness can influence person perception such that greater attraction is positively associated with more favorable judgments.\(^{38,39}\) Given that our primary dependent variables in this study are credibility assessments, before data collection, we matched the relative attractiveness of the experts used as stimuli in this study. Results suggested that attractiveness would not covary with the independent variables (e.g., gender). We included the attractiveness question in the main study as a manipulation check.

**Results**

**Manipulation Check**

**Knowledge**

The knowledge manipulation check indicated that our manipulation of knowledge was successful for each expert; that is, the high-knowledge expert was perceived as more knowledgeable than the low-knowledge expert (\( F(1, 135) = 6.31, p = 0.013 \)) (high knowledge \( M = 39.83, SD = 8.53 \) versus low knowledge \( M = 35.97, SD = 9.33 \)). Because knowledge was rated on a 10-point scale with five items per construct, the possible range in ratings was from 5 to 50. Thus, both experts were rated as relatively knowledgeable. As expected, we found that one expert was significantly less knowledgeable than the other, with a medium effect (\( \eta^2 = 0.042 \)).

**Attractiveness**

To ensure that the potential covariate of attractiveness was independent of the manipulations in this study,\(^{40}\) we matched the female and male experts on attractiveness before the manipulations were tested. We then tested this manipulation check in our study sample by using the following question about each expert: how physically attractive did you find this expert witness on a 10-point Likert scale (not at all attractive (1) to extremely attractive (10)). A significant difference in attractiveness emerged (\( F(1, 135) = 4.65; p = 0.033; \eta^2 = 0.034 \)). The female expert was rated as significantly more attractive (\( M = 5.01, SD = 1.78 \)) than her male counterpart (\( M = 4.31, SD = 2.01 \)) (a small to medium effect). Thus, attractiveness was an unexpected confounder that may diminish some portion of the effect of gender or knowledge or both (the portion associated with attractiveness) on outcomes.

There is a debate in the literature about how to address the confounders in multivariate analysis of covariance (MANCOVA).\(^{41}\) One key determinant on whether analysis of covariance (ANCOVA) can be implemented when a covariate and independent variables are confounded (as in our study) is whether the covariate arises by chance, or it is more likely that a meaningful difference between groups on the covariate is systematically delineated by the independent variable.\(^{42–46}\) MANCOVA is generally appropriate for random assignment designs if the covariate arises by chance because the analysis would be removing only “noise variance from group, not anything substantive about group.” (Ref. 41, p 45).\(^{44}\) In our study, attractiveness is likely to have differed between the male and female experts by chance. We have no reason to believe that the gender of the expert is the factor that influences the difference in attractiveness or that this difference would generalize to all female experts. Although it can be difficult to substantiate causal relationships between a covariate and an independent variable,\(^{47}\) it is widely accepted that attractiveness is a dimension independent of gender.\(^{48,49}\) Men and women vary in attractiveness, and these variables should not be conflated. Thus, our MANCOVA that included attractiveness as a covariate was used in the primary analysis to allow the variance introduced by this unexpected covariate to be reduced.

**Main Analyses**

For our primary analyses, we conducted a MANCOVA with the two independent variables (knowledge condition: high versus low; and expert gender: male versus female) on the dependent variables. We included five dependent variables: three credibility dimensions to examine witness credibility at the facet level (trustworthiness, confidence, and likeability), a continuous sentencing variable, and ratings of the defendant’s likelihood of engaging in
future violence (i.e., agreement with the expert’s opinion). We included expert witness attractiveness as a covariate. Because participant age, gender, or race did not moderate any of the effects in the initial model, we did not include them in our final models.

Tests of multivariate normality, multicollinearity, and homogeneity of variance-covariance matrices revealed no significant results.\(^4\)\(^0\)\(^3\)\(^0\) Because of significant violations of Levene’s test of equality of variance (for both sentencing and likeability), we set a conservative \(\alpha\) level of 0.025 for these outcomes and used Pillai’s trace to examine test statistics.\(^4\)\(^0\) Finally, the sample size requirement for MANCOVA procedures (at least 20 participants per cell\(^4\)\(^9\)) was met, with the sample distributed relatively evenly across conditions. Adjusted means and descriptive information by condition are provided in Table 3.

The MANCOVA results indicated that significant multivariate main effects emerged for the knowledge conditions (Pillai’s trace = 0.146; \(F(6, 127) = 4.34; p < .001; \eta^2_p = 0.146\)). There was no significant main effect of expert gender (Pillai’s trace = 0.79; \(F(6, 127) = 2.19; p = .059; \eta^2_p = 0.79\)), indicating that expert witness gender was not systematically related to any of the dependent variables. The interaction between expert knowledge and gender was not systematically related to any of the dependent variables (Pillai’s trace = .02; \(F(6, 125) = .572; p = .720; \eta^2_p = .02\)).

**Follow-Up Analysis to the MANCOVA**

We initially conducted a discriminant function analysis (DFA) to identify how the dependent variables discriminated the high- versus low-knowledge groups. Essentially, DFA flips the approach to understanding the relationship between knowledge (the independent variable, or IV) and the dependent variables (DVs) used in the MANCOVA. Should a dependent variable explain a portion of the separation between high- and low-knowledge groups (i.e., if the DV can help explain the differences in the IV), it is likely that the significant main effect of the MANCOVA is attributable to the relationship between the IV conditions (knowledge in this case) and the particular dependent variable.\(^5\)\(^1\) In this case, our discriminant analysis revealed one discriminant function that explained 100 percent of the variance: canonical \(R^2 = 0.15\) (small effect size).

The discriminant function showed that the differences in knowledge could be explained in terms of one underlying dimension (Wilks’ lambda = 0.85; \(\chi^2(5) = 21.96; p < .001\)). The correlations between outcomes and the discriminant function\(^5\)\(^1\)–\(^5\)\(^3\) revealed that likeability loaded highly onto the discriminant function (\(r = 0.53\)), followed by sentencing recommendation (\(r = 0.26\)), followed by confidence (\(r = -0.20\)), then by chance of committing future acts of violence (agreement with expert) (\(r = -0.17\)), and finally by the low loading of trustworthiness (\(r = -0.04\)). Indeed, the DFA results indicate that the degree of knowledge is being discriminated between low and high (based on the non-standardized canonical discriminant functions evaluated at group means). Although likeability tended to contribute the most to group separation of high versus low knowledge, the difference between knowledge groups may well be related to sentencing recom-

### Table 3  Means (and Standard Deviations) Defined by Expert Gender and Knowledge

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Manipulation Check</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCStestworthiness</td>
<td>WCScertainty</td>
</tr>
<tr>
<td>High knowledge (n = 72)</td>
<td>36.39 (10.04)</td>
</tr>
<tr>
<td>Male expert (n = 35)</td>
<td>35.97 (9.43)</td>
</tr>
<tr>
<td>Female expert (n = 37)</td>
<td>36.78 (10.71)</td>
</tr>
<tr>
<td>Low knowledge (n = 64)</td>
<td>36.26 (8.27)</td>
</tr>
<tr>
<td>Male expert (n = 32)</td>
<td>34.78 (9.30)</td>
</tr>
<tr>
<td>Female expert (n = 32)</td>
<td>37.63 (7.08)</td>
</tr>
<tr>
<td>Combined knowledge</td>
<td></td>
</tr>
<tr>
<td>Male expert</td>
<td>35.40 (9.32)</td>
</tr>
<tr>
<td>Female expert</td>
<td>37.21 (9.09)</td>
</tr>
</tbody>
</table>

WCS, Witness Credibility Scale factors.

* Jurors’ ratings of the percent chance that the defendant would engage in future violent acts (agreement with the expert’s opinion of a high likelihood).

† Negative scores denote higher likelihood of the death penalty, and positive scores denote higher likelihood of a life sentence without parole.
experts were rated as significantly less likeable (M = 35.62, SD = 8.46) than the less knowledgeable expert (M = 39.58, SD = 8.26). No other significant effects of knowledge on the remaining WCM facets (i.e., confidence or trustworthiness), sentencing recommendations, future violence predictions (agreement with the expert), or gender interactions emerged.

Supplemental Analysis

We conducted these supplemental analyses to examine the variable of attractiveness in more depth. Even though the influence of attractiveness was deemed a nonsystematic covariate in the present study, it is still possible that by entering the covariate into the model, “the covariate will in effect get credit for any relationship of their shared variance [with the independent variable] that is also shared with the dependent variable” (Ref. 41, p 45). The result may be a diminished estimate of the relationship between the gender and the dependent variables. This possibility is of particular concern in the present study because the multivariate main effect of expert gender approaches statistical significance at p = .059 when attractiveness is included in the model.

Attractiveness did in fact exert a significant multivariate main effect in the overall model (Pillai’s trace = .087; F(6, 127) = 2.44; p = .038; η² = .087). When attractiveness was removed from the analysis, a significant multivariate main effect for expert witness gender emerged (Pillai’s trace = .98; F(5, 128) = 2.77; p = .021, η² = .098), and the significant main effect for knowledge condition remained (Pillai’s trace = .15; F(5, 128) = 4.52; p < .001; η² = .15). The interaction between the expert’s knowledge and gender was not systematically related to any of the dependent variables when attractiveness was removed from the model (Pillai’s trace = .02; F(5, 128) = 0.607; p = .694; η² = .02).

These results indicate that a portion of the effect of gender on the dependent variables may in fact have been removed in the main analyses (the portion of the effect that covaried with attractiveness). However, as noted above, this effect is more likely to be explained by gender’s covariate relationship with attractiveness in our particular stimuli. Thus, the implication is that exploring the potential effect of gender (as possibly mediated by attractiveness) was not theoretically supported.

Discussion

In this study, we experimentally manipulated level of knowledge in an expert forensic mental health professional’s testimony on the stand in a mock-trial paradigm. We sought to test the relation between lower and higher degrees of demonstrated expert knowledge and juror perceptions of expert credibility, agreement with the expert, and sentencing decisions. We also tested for potential moderating effects of expert gender. Our knowledge manipulations were successful from an empirical standpoint, operationally defining high versus low demonstrated expert knowledge.

We hypothesized that high knowledge would yield increased credibility as well as increased agreement with the expert. Although knowledge did exert an effect on one facet of credibility (i.e., likeability), it did so in a manner counter to our predictions. Knowledge influenced perceptions of expert likeability such that the expert with lower knowledge was paradoxically perceived as more likeable than the higher knowledge counterpart. The second part of our hypotheses that predicted a positive relationship between knowledge and agreement with the expert was not evidenced in this study. In other words, it appears that our defined levels of very knowledgeable versus less knowledgeable did not influence the mock jurors’ ultimate opinions of the defendant’s risk of future violence (agreement with the expert) or sentencing. Thus, knowledge manipulations influenced perceptions of some facets of credibility, yet carried little predictive utility in understanding mock jurors’ ultimate decisions. Our results do not necessarily imply that an expert’s knowledge has little effect on perceptions of credibility and subsequent juror decisions. Let us examine alternative explanations of our findings.
It is plausible that this study evidenced a ceiling effect, likely to exist in actual testimony, where the peripheral cue of being an expert extended a blanket influence of knowledge. Recall that differences between the low- and high-knowledge experts were statistically significant and yielded a medium effect. However, both experts were perceived as relatively knowledgeable. Moreover, knowledge levels did not contribute to credibility outcomes except in regard to likeability. Mock jurors also did not differentiate between very knowledgeable and less knowledgeable experts for agreement ratings and sentencing. These findings collectively suggest that mock jurors may have relied on the courts’ discretion in allowing only qualified people with specialized knowledge to take the role of expert. That is, jurors may make an assumption that the expert is knowledgeable without critically evaluating the foundation of his or her knowledge. These results align with previous research that suggests the primary persuasive influence in expert testimony is the witness’s status as an expert. Research has shown that jurors may not sufficiently evaluate the foundational research of expert opinions and that they may defer to the clinical opinion of the expert over an opinion rooted in actuarial evidence. The current study adds to the literature. Even when knowledge is varied (high versus low), there does not appear to be a critical evaluation of the witness, perhaps due to the witness’s qualification as an expert. Thus, differential decision-making that could otherwise result from differences in expert knowledge may not be elicited.

However, low knowledge did increase the expert’s likeability, and that result suggests that additional social-cognitive processes are at work. The negative relation that emerged between level of expert knowledge and perceived likeability implies that aspects of higher versus lower knowledge may influence expert likeability. Although we are cautious about speculating on underlying processes that were not directly examined in this study, it is possible, for example, that learning about an expert’s qualifications would create a psychological distance between the expert and the mock juror. Social psychology research supports the competence–liking paradox; that is, the person with the most knowledge is often not the most liked. In court, and in life, however, it would seem beneficial to like the more knowledgeable person, as he may increase our chances of being correct and competent. Nevertheless, likeability for a knowledgeable expert comes at a cost to the juror, who may feel that he pales in comparison to the all-knowing expert. Perceptions of similarity and mutual liking decrease when a person perceived as superior is a factor. In fact, the pratfall effect suggests that competence with some degree of fallibility is perhaps the most liked combination and that a juror’s gender and self-esteem may play into this phenomenon. Another possibility is that the highly knowledgeable experts were disliked because of character cues elicited from the high-knowledge content (e.g., perceived narcissism). Thus, it is plausible that differences in knowledge (e.g., the perceived narcissism in very knowledgeable experts) are more or less interpreted as differences in likeability (e.g., less likeable). In other words, experts may benefit from Baldoni’s recommendation: “Never act like the smartest guy in the room.”

When it comes to credibility, mock jurors may defer to the court and view very knowledgeable and less knowledgeable experts as knowledgeable because of their expert status. Thus, although it may seem that differences in knowledge have little influence on credibility determinations, differences in demonstrations of knowledge (e.g., high- versus low-knowledge presentations on the stand) may elicit psychological and peripheral cues to an expert’s likeability. The evaluative, social, and cognitive influences that could be responsible for the negative knowledge–likeability link found in this and other research deserve future empirical attention, particularly given the potential influence of expert likeability on mock juror decision-making.

Overall, the degree to which jurors are sensitive to differences in an expert’s knowledge is not clear. Perhaps a continuum of perceived knowledge exists and exerts a meaningful influence on credibility. More likely, however, jurors assign a knowledge threshold to the person who is deemed an expert by the courts, consistent with heuristic models of jurors’ evidence interpretation. Thus, perhaps the relative quality of the witness’s expertise lacks a significant, observable influence on decision-making. This finding dovetails with prior witness credibility research. Despite the influence of manipulations on overall credibility, the components of credibility often lack direct or explicitly observable influence on individual jurors’ explicit decision-making.

The finding that differences in knowledge may affect the expert’s perceived credibility but that ex-
Expert Witness Knowledge

pert differences did not translate into differences in jurors’ ultimate decisions is potentially good news. The decision of the trier of fact is supposed to be based on the content of testimony, the substance of a case, and the strength of evidence.65 These findings add to the body of research showing that other variables affect the decision-maker, but only incrementally. That is, a variable such as expert witness knowledge is but one of many pieces of information decision-makers must integrate in formulating a decision. Experts and trial consultants may still benefit from recognizing that in close cases (i.e., those in which the verdict could go either way) or in cases in which opposing experts testify, expert knowledge may exert a substantive influence. In such instances, it might be beneficial to keep in mind that displays of knowledge may not always work in one’s favor, at least to the extent that they diminish one’s likeability.17

Effects of Expert Witness Gender

We found that expert gender had no effect on perceptions of credibility or mock juror decisions. Further, no statistically significant interactions regarding expert gender emerged in the present study. These results are encouraging: they suggest that jurors may not be using gender as a peripheral cue to assess expert knowledge or credibility.

Implications for Testifying Experts and the Attorneys Who Select Them

We constructed large differences in high- and low-knowledgeable experts in this study. However, the mock jurors did not pick up on the differences in the experts to the degree that we expected. These findings suggest that in uncontested cases or cases where the evidence is overwhelmingly strong for one side, the expert’s basic credentials, accomplishments, and demonstrated knowledge may not make much of a difference to jurors. Experts and attorneys in such cases may not have to fret about relatively unaccomplished experts; so long as they meet a threshold level of perceived knowledge, various credentials may not matter. For example, it may make no difference whether the expert attended an Ivy League university or a lesser known institution, is board certified, or has published in scientific journals.

What this research cannot speak to is whether differences in the level of experts’ knowledge would make a difference to judges, for example, in bench trials. Judges are probably more sophisticated about discerning relative degrees of expert knowledge. Furthermore, because our participants were exposed to only one expert, the results cannot show whether judges or jurors would notice relative differences in experts’ knowledge if there were opposing experts in a single case. Whereas a meta-analysis found similar effects of unopposed and opposed expert testimony on juror decision-making,66 other studies have revealed particular contexts in which opposing testimony may have a uniquely strengthened effect.13 Perhaps if the high- and low-knowledge experts had been compared side by side, their differences would have become more salient, and a stronger effect would have been found.

Strengths, Limitations, and Future Directions

The knowledge manipulations used in this study were developed by amalgamating conceptual components from a variety of prior research projects. Ours is the second study to use these knowledge manipulations (see Ref. 10, for the first use). The current study was the first to test the unique effects of knowledge on mock jurors’ determinations of witness credibility and decision-making. A strength of this design is the resultant ability to interpret direct causal relations of expert knowledge and gender to credibility and case-related decision-making. Additional strengths adding to ecological validity were using actual PhD forensic psychologists as experts in the video-taped scenarios and filming the stimulus videos in a well-simulated environment.

However, to achieve the control needed for experimental manipulation of expert knowledge in this preliminary study, we did not fully capture some real-world elements of a capital trial. Limitations include the lack of voir dire or deliberation and the use of a college student mock juror sample.67 College students often provide a large, easily accessible population for the purposes of initial mock jury research.68 While a review of jury simulation research concluded that the use of students as mock jurors is not necessarily a cause for concern,69 recent research suggests some differences between college and community samples.70,71 Nevertheless, the use of a college sample has been deemed no more problematic to generalizability than other common variables (e.g., trial context and jurisdiction).68

Other limitations of our sample include that it was largely Caucasian (79%), female (81%), and young (average age, 19). Although the characteristics of this
sample do not reflect that of an average jury pool, the mock jurors in this study were jury-eligible citizens and may serve in actual trials at some point. A replication of our results with a more diverse sample within a paradigm that further extends the realistic nature of the trial process would allay some validity concerns, increase the generalizability of the findings, and increase the confidence that the field can place in these results. Analyses were also complicated because the particular female expert in this study was perceived as more attractive than the particular male expert, suggesting that our MANCOVA results should be interpreted with caution. It is possible that the gender manipulation was somewhat weakened by the difference in attractiveness and thus underestimated the relationship between gender and the dependent variables. In future research, investigators should seek to avoid confounding due to attractiveness, possibly by including multiple male and female experts for comparison.

Studying witness knowledge in a capital proceeding potentially limits the generalizability of our findings to other court proceedings. Of course, this criticism is not unique to capital proceedings. The same argument could be made for any other potential proceeding. Had we chosen to study expert knowledge in a civil commitment proceeding, for example, those findings might have been relevant only for other civil commitment proceedings. We chose to study expert knowledge in a capital case for several reasons. First, because the possibility of a sentence of death makes the verdict different from that which would result in other sentences and capital trials are among the most contentious cases, mock jurors’ motivation to attend to the task and to the expert may have been maximized in this context. Second, lawyers and experts may seek the most consultation, given the resources devoted to a capital case and the high stakes (death versus life) partially contingent on testimony effectiveness.

Third, in all the other studies we have published that involved experimentally manipulating elements of expert credibility, we have used the same basic mock-trial stimuli (Table 1). For meaningful comparison of the findings from the current study with the body of research that has developed on witness credibility, we wanted to hold constant as many details as possible, other than the credibility behaviors that have been manipulated across the various studies. Finally, most prior research on expert witness testimony in which clinical versus actuarial testimony effectiveness was varied (related to expert witness knowledge) has been in capital sentencing paradigms, which allows us to build on this line of research.

We also note that the witness’s self-proclaimed expertise (e.g., the expert’s statement that “as far as I know I’ve never been wrong”) may have introduced a confounder of perceived arrogance coupled with high knowledge. To the extent that this confounder was present and systematically affected perceived likeability, this aspect of the knowledge presentation may have influenced more than just perceived expertise by lessening expert likeability and hampering the effects of increased confidence in the highly knowledgeable expert. Future research should explore the relative influence of various types of high expert knowledge displays on the stand.

Overall, manipulations of expert knowledge did not affect credibility or significantly predict mock jurors’ decisions in the hypothesized manner. In this discussion, we presented hypotheses about why these findings may have emerged, emphasizing support found for heuristic models’ explanatory value in understanding how expert testimony may influence jurors’ evaluation of an expert’s credibility. Given the centrality of expert knowledge to the courts’ reliance on expert testimony, future research should seek to clarify its role in juror evaluations of expert evidence. In short, in answer to the questions we originally set out to explore, it appears that mock jurors do notice variations in expert witness knowledge; however, this difference may not carry weight when it comes down to influences on evidence interpretation and decision-making.

Acknowledgments

The authors thank Michael and Desiree Griffin for their assistance in preparing the video stimuli and Maddy Semon for data collection and management.

References

Expert Witness Knowledge

12. Cooper J, Neuhaus JH: The “hired gun” effect: assessing the effect of pay, frequency of testifying, and credentials on the perception of expert testimony. Law & Hum Behav 24,149–71, 2000
60. Deaux K: To err is humanizing: but sex makes a difference. Represent Soc Psychol 3:20–8, 1972