

Warmth and Competence on the Witness Stand: Implications for the Credibility of Male and Female Expert Witnesses

Tess M. S. Neal, MA, Rosanna E. Guadagno, PhD, Cassie A. Eno, PhD, and Stanley L. Brodsky, PhD

In this study, we examined how manipulations of likeability and knowledge affected mock jurors' perceptions of female and male expert witness credibility ($n = 290$). Our findings extend the person-perception literature by demonstrating how warmth and competence overlap with existing conceptions of likeability and knowledge in the psycholegal domain. We found that experts high in likeability, knowledge, or both were perceived equally positively, regardless of gender, in a death penalty sentencing context. Gender differences emerged when the expert was low in likeability or knowledge. In these conditions the male expert was perceived more positively than the comparable female expert. Although intermediate judgments (e.g., perceptions of credibility) were affected by our manipulations, ultimate decisions (e.g., sentencing) were not. Implications for theory and practice are discussed.

J Am Acad Psychiatry Law 40:488–97, 2012

Women and men are perceived differently in many situations,^{1–4} including when they serve as expert witnesses in legal settings.^{5–7} Individuals may testify in court as expert witnesses if they have specialized knowledge that will assist the trier of fact in determining the relevant legal issue.⁸ Attorneys often rely on the testimony of an expert witness as part of their trial strategy to present their case to the trier of fact. The implications of differential perceptions of expert witnesses on the basis of the expert's gender are important to understand in legal settings, as juror decision-making may be influenced by perceptions of expert witnesses. In this study, we examined how

male and female experts are perceived when aspects of credibility, knowledge (e.g., competence), and likeability (e.g., warmth) are manipulated.

Credibility has been discussed in many domains and is understood to be an important aspect of person perception, accounting for up to 82 percent of variance in global impressions.⁹ Although researchers have studied the construct from a variety of disciplines, we focus on social perception and psycholegal research, which appear to be most germane to the present research. The domains of competence and warmth are thought to drive stereotypes in the social perception literature.^{10,11} Researchers from this perspective argue that we initially categorize individuals as being high or low in both of these domains. The stereotype content model^{10,11} differentiates stereotyped groups along two dimensions: competence and warmth. Competence is driven by perceptions of confidence, skillfulness, and capability, whereas warmth is driven by perceptions of friendliness, good-naturedness, and sincerity.¹²

The domains of competence and warmth overlap conceptually with the knowledge and likeability sub-components of witness credibility theory proposed

Ms. Neal is a predoctoral psychology intern, Department of Psychiatry, University of Massachusetts Medical School, Worcester, MA. Dr. Guadagno is Program Director, National Science Foundation, Washington, DC. Dr. Eno is Adjunct Professor and Department Chair, Department of Psychology, Waldorf College, Forest City, IA. Dr. Brodsky is Professor of Psychology and Coordinator of the Psychology-Law PhD Concentration, Department of Psychology, University of Alabama, Tuscaloosa, AL. This research was presented at the Fourth International Congress on Psychology and Law, March 2011, Miami, FL. Address correspondence to: Tess M. S. Neal, MA, Department of Psychiatry, University of Massachusetts Medical School, WSH 8C, 55 Lake Avenue North, Worcester, MA 01655. E-mail: tmneal@crimson.ua.edu.

Disclosures of financial or other potential conflicts of interest: None.

by Brodsky *et al.*¹³ They conceptualized four domains of witness credibility: knowledge, likeability, trustworthiness, and confidence. Relevant to the present investigation, conceptions of competence and warmth converge with two domains in this model: knowledge (i.e., competence) and likeability (i.e., warmth).

Warmth and Competence Stereotypes

According to the stereotype content model (SCM), the dimensions of warmth and competence result in four different patterns of stereotypes based on combinations of warmth (high/low) and competence (high/low).^{10,11} People perceived as both warm and competent elicit uniformly positive emotions and behavior, whereas people perceived as lacking warmth and competence elicit uniform negativity.¹⁴ People perceived as high on one dimension but low on the other elicit reliably ambivalent affective and behavioral responses.¹⁴ For example, people perceived as high on competence but low on warmth elicit envy and competitive behavior, and people low on competence but high on warmth elicit pity and neglect.^{14,15} The SCM has been validated in several different countries and cultures; in fact, its authors present it as a pancultural tool for predicting group stereotypes.¹²

Gender Role Expectations and Perceptions

According to social role theory, normative expectations for behavior differ for men and women.¹⁶ Women are generally expected to be more warm and communal than men. They are thought to be more emotionally expressive, interpersonally sensitive, and concerned about others. Men are generally expected to be more competent and agentic than women. For instance, men are expected to be more controlling, independent, and assertive.^{3,16,17}

Research combining social role theory with the SCM has found that women risk being negatively perceived more than men if social roles are violated. Few studies to date have revealed contexts in which women are perceived as high in both warmth and competence, possibly due to the restrictions of social roles. Women essentially have the choice to be perceived as warm but incompetent^{18–22} or as competent but cold.^{18,21–23} However, many men are perceived as both competent and warm. For example,

working men who become fathers maintain perceived competence, but also gain perceived warmth, reaping the social benefits of eliciting universal positive reactions from others.^{14,17} Working women who become mothers do not gain perceived warmth after having children and in fact lose perceived competence.¹⁷

Although women perceived as high in both competence and warmth appear to be rare, Guadagno *et al.* (manuscript submitted) found one situation in which women are rated as high in both domains: women successful in politics who are perceived as nurturing mothers with supportive husbands (e.g., women like those represented by Sarah Palin's public persona, the hockey mom). Across time and geographic region, these authors found that the hockey mom was perceived as warm and competent, but only when she was presented as a politician. Thus, these findings may indicate that women actually may have more pathways than men do to being regarded as credible. Women can potentially be successful in three of the four combinations of warmth and competence (e.g., cold/competent, warm/incompetent, and warm/competent). It may be the case that men can be successful in only two of the four combinations (e.g., cold/competent and warm/competent). That is, men who are viewed as incompetent may stand little chance to be viewed as credible, whereas women who are seen as incompetent may still have a chance to be regarded as credible. One goal of the present investigation was to see how these findings generalize to women in the courtroom by assessing the factors that affect perceptions of credibility among female and male expert witnesses.

One recent study conducted on expert witnesses⁶ provides some support for these assertions. The researchers compared the credibility of male and female expert witnesses while manipulating levels of eye contact the witnesses made with the attorney and jury. The results indicated that female experts were credible regardless of their eye contact level, whereas men were credible only if they maintained assertive (high) eye contact. These findings suggest men must be regarded as competent to enhance perceptions of credibility, but that women might be thought of as credible when they are competent or warm.

The Current Study

The purpose of the current study was to extend the literature by examining how female and male experts

are perceived when their likeability (i.e., warmth) and knowledge (i.e., competence) are manipulated. We use the terms likeability and knowledge because they have been established in the mental health-law literature^{6,13,24,25}; however, the evidence just reviewed suggests that they overlap conceptually with the warmth and competence domains in the social psychology literature and we present them as related constructs.

Likeability

We defined likeability as the degree to which an expert is friendly, respectful, kind, well mannered, and pleasant.^{13,25} This operational definition was drawn from a variety of sources and focused on behaviors that could be manipulated in the context of testimony. These were our specific, manipulated conceptions of high and low likeability:

High Likeability

Highly likeable behavior consisted of the following components: consistent use of we or us when discussing members of the scientific community or humanity as a whole,²⁶ moderate levels of smiling,²⁷ modest statements and conclusions (e.g., relatively certain or, “we do not know everything there is to know in psychology”),²⁸ consistent eye contact with lawyer and jury,²⁹ informal speech (i.e., little use of technical jargon and use of names),³⁰ and a self-effacing presentation style.³¹

Low Likeability

The behavior of experts with low likeability was made up of these elements: no use of we or us (e.g., psychologists, or people, to speak about people in general), no smiling, excessive statements of certainty of conclusions/arrogance, inconsistent eye contact, highly technical jargon and frequent formal references (e.g., the client, the defendant), and a narcissistic presentation style.

Knowledge

We defined knowledgeable 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. To manipulate this variable in the study, we operationally defined knowledge by drawing on literature from a variety of sources to compile a list of behaviors associated with ratings of knowledge. Our literature review identified the following elements associated

with high knowledge: degree of assertiveness,³² substantive content and clarity of testimony,^{33,34} credentials,^{8,33,35} relevant experience,^{8,35,36} self-proclaimed expertise,³⁷ and familiarity with the case.³³ We also manipulated perceptions of high and low knowledge.

High Knowledge

High levels of knowledge were demonstrated by strong educational credentials (e.g., educated at Yale, ABPP [American Board of Professional Psychology] certified in Forensic Psychology, and a history of academic publication in this area of expertise),^{8,33,35} solid relevant experience (e.g., risk assessment researcher who has conducted over 100 such assessments over 14 years),^{8,35,36} consistent clarity and substantive content of communication,^{33,34} moderate assertiveness (e.g., “as far as I know I’ve never been wrong” when queried about awareness of clinician error),³² self-proclaimed expertise (e.g., “In my expert opinion . . .”),³⁷ and familiarity with the case (e.g., multiple interviews with the defendant).³³

Low Knowledge

The variable of low knowledge consisted of the following behaviors: no mention of educational credentials, minimal relevant experience (e.g., two years as a psychotherapist, no previous experience in risk assessment), inconsistent clarity and substantive content of communication, a low level of assertiveness (e.g., no, when queried about awareness of clinician error), no self-proclaimed expertise, and inadequate familiarity with the case (e.g., one short interview with the defendant the week the case went to trial).

Hypotheses

In light of the existing literature, we hypothesized the following:

The most credible experts and those eliciting the highest agreement ratings from mock jurors would be experts high in likeability and knowledge.

The least credible experts with the lowest agreement ratings would be those low in likeability and knowledge.

In the high-likeability condition, female experts would be perceived as more credible and have higher agreement ratings than would male experts, but in the low-likeability condition male experts would be regarded as more credible with

higher agreement ratings than would female experts.

No interaction between gender and the knowledge of the expert.

Female experts high in likeability but low in knowledge would be viewed as equally credible and elicit equal agreement ratings when compared with male experts. Although SCM would predict women would be more credible than men with this combination of likeability and knowledge,¹⁷ social role theory would predict that the man would be perceived as more competent, due to the perception that the role of the expert is a masculine one. Therefore, we predicted that both theories would be true, washing out any difference for this particular combination.

Males and females high in knowledge but low in likeability would be regarded as equally credible and would elicit equal agreement ratings.

Method

Participants

Undergraduate students ($n = 265$) at a large public university in Alabama were recruited through the Psychology Subject Pool and participated for course credit. The gender composition of our sample was almost evenly split (51% female) and ranged in age from 18 to 36 years (mean (M) 19.64; SD 2.01). Seventy-six percent were Caucasian, 16 percent were African-American, and 8 percent were from other ethnic backgrounds. The Supreme Court decided in *Witherspoon v. Illinois*³⁸ that jurors in capital murder trials must be death qualified; that is, they must be willing and able to consider capital punishment. Because our stimulus material was based on the sentencing phase of a capital murder trial, those individuals who indicated absolute opposition to the death penalty were not included in our analyses ($n = 25$, reducing the total sample size from 290 to 265). Data from participants ineligible based on the *Witherspoon* criteria were equally distributed across our design, and their data did not significantly differ from those of the eligible participants on any demographic variable in our dataset. It should be noted that Alabama, where the participants for this study were sampled, is a state in which defendants can be sentenced to death if convicted of a capital crime.

Stimuli

This study was a 2 (male vs. female expert) \times 2 (high vs. low likeability) \times 2 (high vs. low knowledge) between-subjects factorial design. We developed eight videos, each approximately seven minutes long, to match the eight conditions of the study. The script for the videos was adapted from an actual jury sentencing proceeding described by Krauss and Sales.³⁹ Previous research has successfully used the same basic script to examine expert witness credibility.^{6,24,25} The script portrays a forensic expert witness testifying about his or her evaluation of a convicted murderer and about the defendant's likelihood of committing future violent acts. The expert testifies under both direct and cross-examination. For each condition, knowledge and likeability of the witness were manipulated. One male expert and one female expert matched for age, race, clothing, and attractiveness were filmed in each condition.

Materials

Witness Credibility Scale

The Witness Credibility Scale (WCS) was used to assess the credibility of the expert.¹³ The scale contains 20 bipolar adjectives on a 10-point Likert scale. Higher scores indicate greater credibility ratings. A few examples of these bipolar adjectives include: unkind (1) to kind (10); dishonest (1) to honest (10); and shaken (1) to poised (10). Factor analysis identified four independent domains: trustworthiness, confidence, likeability, and knowledge.¹³ Alpha coefficients were reported for each domain as follows: confidence (.88), likeability (.86), trustworthiness (.93), knowledge (.86), and overall credibility (.95).

Sentencing Ratings

Participants were asked to write down a percentage (1%–100%) indicating how likely they thought the defendant was to commit future acts of violence. The expert in the video testified about the substantial likelihood that the defendant would reoffend, and so this question allowed us to assess the substantive agreement of the participant with the expert witness; in effect, how believable the expert was. Participants were also asked to rate on Likert-type scales how likely they would be to sentence the defendant to the death penalty or to life in prison without the possibility of parole (LWOP). These are the only two sentencing options available to defendants found guilty of capital murder.

Demographics

A basic demographic questionnaire eliciting the participant's gender, age, and level of agreement with the death penalty was included to assess possible differences between subject groups.

Manipulation Check

A manipulation check allowed us to assess the strength of the manipulations. Participants rated three questions on 10-point Likert-type scales: How likeable did you find this expert witness? (not at all likeable (1) to extremely likeable (10)); how knowledgeable did you find this expert witness? (not at all knowledgeable (1) to extremely knowledgeable (10)); and how physically attractive did you find this expert witness? (not at all attractive (1) to extremely attractive (10)). We controlled for the biasing effect of attractiveness because studies have shown that attractive people are judged more favorably than unattractive people.^{40,41}

Procedure

As per the approved institutional review board protocol for this study (approval was granted by the University of Alabama Institutional Review Board), participants were provided with information about the study procedures and provided informed consent before viewing a randomly assigned video condition. After watching the video, they individually completed the questionnaires, including the Witness Credibility Scale,¹³ a sentencing rating form that included both a sentencing decision and a rating (0%–100%) of substantive agreement with the experts' testimony, basic demographics, and the manipulation check. At the completion of the study, the participants were debriefed.

Results

Pilot Study

We conducted a between-subjects multivariate analysis of variance (MANOVA) to compare mock juror ratings ($n = 22$) of still photographic images of each expert witness, to ensure that the experts were matched on credibility, likeability, knowledge, attractiveness, jurors' substantive agreement with the expert, and juror's recommended sentence, before commencing data collection with the videos. Results indicated that there were no significant differences based on the images of the male and female experts

(Wilks' lambda = 0.73, $F(8,14) = 0.65$, $p = .73$, $\eta_p^2 = 0.27$).

Primary Analyses

Manipulation Check

Our manipulation check items indicated that our manipulations of knowledge and likeability were successful for each expert; that is, the knowledge manipulation worked ($F(1,262) = 25.79$; $p < .001$; high knowledge ($M = 86.92$, $SD = 14.58$) vs. low knowledge ($M = 75.26$, $SD = 21.40$)), as did the likeability manipulation ($F(1,262) = 226.76$; $p < .001$; high likeability ($M = 72.24$, $SD = 25.49$) versus low likeability ($M = 23.34$, $SD = 27.18$)). Further support indicating that we successfully manipulated these individual constructs is that the likeability manipulation did not affect the knowledge ratings ($F(1,262) = 2.19$; $p = .14$) and the knowledge manipulation did not affect the likeability ratings ($F(1,262) = .03$; $p = .87$).

Main Analyses

For our primary analyses, we conducted a MANOVA with our three independent variables (gender: male vs. female expert; likeability condition: high vs. low likeability; and knowledge condition: high vs. low knowledge) on our collection of dependent variables (WCS score, rating of substantive agreement with the expert, and a continuous sentencing variable). In the initial model, we examined whether participant age, gender, or race moderated any of our effects. They did not, and so we did not include them in our final models. Significant multivariate main effects emerged for knowledge condition (Wilks' lambda = 0.95; $F(3,231) = 4.05$; $p = .008$, $\eta_p^2 = 0.05$) and likeability condition (Wilks' lambda = 0.71; $F(3,231) = 31.71$; $p < .001$; $\eta_p^2 = 0.29$). These multivariate findings indicate that the knowledge and likeability manipulations were each significantly related to at least one of the dependent variables, which we explore further later. The main effect of gender of expert was not significant (Wilks' Lambda = 0.97; $F(3,231) = 2.07$; $p = .11$, $\eta_p^2 = 0.03$) indicating that expert gender was not systematically related to any of our dependent variables.

We then conducted targeted univariate planned comparisons to test our *a priori* hypotheses. Our first hypothesis, that the most credible experts would be those high in both likeability (e.g., warmth) and knowledge (e.g., competence) was supported

($F(1,233) = 3.69; p = .05; \eta_p^2 = 0.02$). The expert who was highly likeable and highly knowledgeable was rated as significantly more credible ($M = 8.14, SD = 1.04$) than any of the other three combinations of likeability and knowledge (range of the mean = 6.22–7.71, range of the SD = 1.02–1.28). Although we expected that the pattern of results for participant ratings of testimony agreement would parallel the credibility result findings, no significant differences emerged in agreement ratings based on the likeability or the knowledge of the experts.

Our second hypothesis, that the least credible experts would be those low in both likeability and knowledge, was also supported ($F(1,233) = 9.29; p = .003; \eta_p^2 = 0.04$). The expert who was both low in knowledge and likeability was rated as significantly less credible ($M = 6.22, SD = 1.28$) than any of the other three combinations of likeability and knowledge (range of the mean = 6.81–8.14; range of the SD = 1.02–1.27). As with the first hypothesis, no significant differences emerged in agreement ratings based on likeability or knowledge of the experts.

We found partial support for our third hypothesis, which predicted that the likeability of the expert would differentially affect credibility ratings of the male and female experts and ratings of agreement with expert testimony. In the low likeability condition, the female expert was rated as less credible ($M = 6.25, SD = 1.21$) than the male expert ($M = 6.69; SD = 1.18; F(1,233) = 4.96; p = .027; \eta_p^2 = 0.02$), and the participants agreed more with the male expert ($M = 73.37, SD = 17.50$) than with the female expert ($M = 66.31, SD = 15.43; F(1,233) = 5.92, p = .016, \eta_p^2 = .03$). Although we expected that the female expert would be regarded as more credible than the male expert in the highly likable condition, there were no differences in credibility ratings ($F(1,233) = 0.84; p = 0.36; \eta_p^2 = 0.02$), nor were there differences in participant agreement ratings ($F(1,233) = 2.93; p = .09; \eta_p^2 = .01$).

The fourth hypothesis predicted no interaction between gender of expert and knowledge. According to the MANOVA results just described, this hypothesis was supported: no significant effects were found for our collection of dependent variables (Wilks' lambda = 0.99; $F(3,231) = .74; p = .53; \eta_p^2 = 0.01$). However, we noted an interesting pattern in the univariate results, which we describe later, in the Exploratory Findings section.

Our fifth hypothesis predicted no differences in credibility ratings or ratings of agreement for the male and female experts in a three-way interaction between expert gender, low knowledge, and high likeability. This hypothesis was supported, as no differences emerged in credibility ratings between the male expert ($M = 7.92, SD = 1.19$) and the female expert ($M = 7.52, SD = 1.32; F(1,233) = 1.68; p = .20; \eta_p^2 = .07$), and no differences emerged in percentage of agreement with the male expert ($M = 76.12, SD = 16.37$) and female expert ($M = 70.34; SD = 17.34; F(1,233) = 1.47; p = .23; \eta_p^2 = .01$).

We predicted that highly knowledgeable but unlikeable experts, whether they were men or women, would be regarded as equally credible and would elicit similar ratings of substantive agreement. The data support this hypothesis for credibility ratings ($F(1,233) = 1.01; p = .32; \eta_p^2 = 0.01$), where no differences emerged in credibility ratings of the male expert ($M = 6.96, SD = 1.02$) and female expert ($M = 6.66, SD = 1.04$). The data also support the hypothesis for agreement ($F(1,233) = 2.04; p = .16; \eta_p^2 = 0.01$), with no differences in agreement with the male ($M = 73.79, SD = 15.39$) versus the female ($M = 67.68, SD = 15.37$) experts. An interesting note is that the woman was never rated as more credible than the man in any analysis, and under no conditions did participants indicate more agreement with the female expert than the male.

Exploratory Analyses

We explored how the gender of the expert and manipulations of likeability and knowledge affected a continuous sentencing rating. Because of the sparse literature on this topic, we treated this query as an exploratory research question. We included the sentencing decision as a dependent variable in the factorial MANOVA, when conducting the *post hoc* analyses. Results did not reveal any significant main effect or interaction for this outcome variable.

Although we made a few *a priori* predictions about the interactions between expert gender, knowledge, and likeability, we did not formulate many specific hypotheses about how the male and female experts would differ in the low-knowledge and -likeability conditions. Our exploratory analyses revealed a consistent pattern of results uniformly in favor of men within these conditions (i.e., low knowledge and low likeability; see Fig. 1). For instance, participants rated the low-knowledgeable male expert as signifi-

Warmth and Competence on the Witness Stand

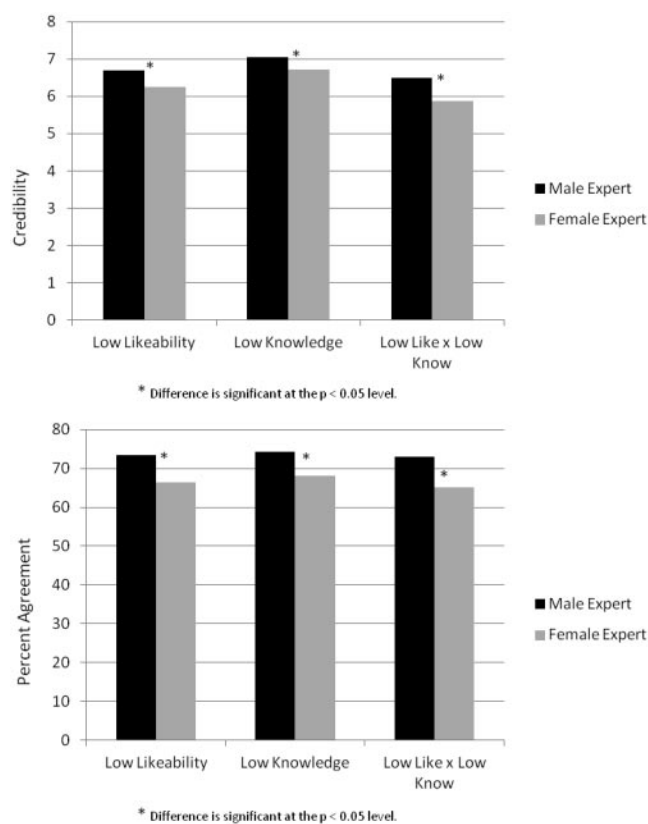


Figure 1. Interaction between the expert's gender, knowledge, and likeability on credibility ratings and percentage agreement with the expert.

cantly more credible ($M = 7.06$, $SD = 1.41$) than the comparable female ($M = 6.71$, $SD = 1.53$; $F(1,233) = 5.94$; $p = .016$; $\eta_p^2 = .03$) and participants also agreed more with the substantive content of the low knowledgeable male's testimony ($M = 74.28$, $SD = 18.0$) than with the comparable female's ($M = 68.02$, $SD = 16.66$; $F(1,233) = 5.16$; $p = .024$; $\eta_p^2 = .02$). The same pattern emerged for the experts low in likeability: participants rated the male expert as more credible than the female and, consistent with Hypothesis 3, agreed more with him. Further, when the expert was both low in likeability and knowledge, participants found the male expert significantly more credible ($M = 6.49$, $SD = 1.26$) than the female expert ($M = 5.87$, $SD = 1.25$; $F(1,233) = 4.83$; $p = .029$; $\eta_p^2 = .02$). Likewise, the participants agreed significantly more with the substantive content of the low likeability and low in knowledge male expert ($M = 73.05$, $SD = 19.11$) compared with the corresponding female expert ($M = 65.03$, $SD = 15.65$; $F(1,233) = 4.17$; $p = .042$; $\eta_p^2 = .02$; see Fig. 1).

While conducting our data analyses, we noticed that the participants' credibility evaluations had a larger amount of variability between the low ($M = 6.49$, $SD = 1.21$)- and high ($M = 7.93$, $SD = 1.17$)-likeability conditions than between the low ($M = 6.89$, $SD = 1.47$)- and high ($M = 7.48$, $SD = 1.27$)-knowledge conditions. Based on the differences observed in these ranges, we decided to explore how likeability and knowledge would each uniquely account for the total credibility score, as well as how much variability they could account for together in the total credibility score. To explore this question, we conducted a multiple regression between the likeability manipulation check rating, the knowledge manipulation check rating, and the total credibility score.

The results indicated that the full $R^2 = .602$, indicating that the model (i.e., regressing the likeability and knowledge ratings on total credibility) accounted for 60.2 percent of the variability in total credibility score. Thus, likeability and knowledge jointly explained a significant amount of variability in the total credibility score ($F(2,238) = 180.30$; $p < .001$). To explore the unique ability of each predictor, we performed a stepwise regression, in which likeability rating was entered in the first step and was joined by knowledge in the second step. The R^2 value for likeability alone was 0.509, a substantial portion of variance in the full model. When knowledge was added in, the R^2 change was small ($\Delta = 0.093$) but significant ($F(1,238) = 55.90$, $p < .001$). These results suggested that while likeability accounted for most of the variability in judgments of expert witness credibility, knowledge added a small but unique ability to explaining the variance in credibility ratings.

Discussion

Previous work has identified warmth and competence as factors accounting for a large portion of the variance in perceptions of others.⁹⁻¹³ Consistent with these reports, we found that experts high in likeability and knowledge were viewed as the most credible, whereas those experts low on these dimensions were seen as the least credible. This finding replicates the person-perception literature and extends it by showing how warmth and competence overlap with likeability and credibility in the psychological domain.

With respect to gender differences in evaluations of our expert witnesses, our results revealed that par-

ticipants evaluated men and women differently when in the role of an expert and when appearing low in likeability or knowledge. The primary implication from this study is that gender of an expert witness matters, but only when the expert is not both high in likeability (e.g., warmth) and knowledge (e.g., competence). In general, our results revealed that women fared more poorly than men, consistent with the literature. We take a social role theory perspective¹⁶ on these findings and suggest that this was the case because the women were in a masculine occupational role, that of an expert witness, and violated normative expectations for likeability. Further, the domain of this case may have been masculine, in that the expert testified about violent recidivism. Had the domain of the case been more stereotypically feminine (e.g., perhaps child abuse), the pattern of results might have differed.^{42,43}

This rationale is further supported by the fact that our pilot study found that without the occupational domain, participants evaluated the experts similarly. That we included only one domain of testimony, that of violent recidivism in a capital murder sentencing context, is a limitation of this study. The SCM^{14,17} predicts that women must be viewed as both warm and competent to compare favorably to men. Perhaps future research can provide more opportunities for male and female experts to demonstrate credible warmth and competence by allowing them to testify in cases with more stereotypical feminine and masculine domains.

The range of credibility assessments for expert knowledge in this study was relatively restricted, which suggests that the participants accepted witnesses as experts and perceived of them as knowledgeable, even when their credentials and relevant experience were not objectively impressive. Although jurors may have had difficulty critically evaluating an expert's knowledge, likeability is a peripheral cue, and therefore easier to evaluate with minimal information processing.^{44,45} The notion that jurors process expert knowledge and likeability information using potentially different processes within the dual-process model of persuasion (i.e., peripheral and central routes) is speculative on our part; still, the data support this interpretation. Future research on perceptions of expert witnesses should include specific measures such as cognitive responses⁴⁶ that provide more direct evidence on this finding.

Our second hypothesis, that the least credible experts would be those low in both likeability and knowledge, was supported; however, the expected parallel pattern of results for participant ratings of testimony agreement did not emerge. Perhaps instead of using peripheral processing, the mock jurors used central processing in their ultimate task of evaluating the evidence and deciding how much they agreed with the experts.⁴⁶ Alternatively, experts low in knowledge and likeability may have still reached the minimum threshold for the mock jurors to evaluate them as experts. This pattern may have been particularly likely, given that the expert witnesses were in their mid-50s and may have received a boost in perceived credibility owing to age. If this was the case, it would follow that differences in substantive agreement may not be affected by manipulations of peripheral cues such as likeability and knowledge. These possibilities suggest avenues for further research to explore why manipulations of likeability and knowledge may affect credibility ratings but not affect substantive agreement with expert testimony.

Although intermediate decisions yielded the expected results (i.e., credibility ratings and substantive agreement), when it came to the ultimate decision a juror must make (the sentence, in this case), no significant differences were found. These findings suggest that although stereotypes of men and women may influence intermediate judgments, ultimate decisions may not be influenced by such stereotypic cues. Similar patterns of results have been found in other studies: although successful manipulations may affect verdict or credibility ratings, one does not always translate into the other.^{47,48} It is possible that, in a different case context with more salient gender cues, ultimate decisions could be affected as well. For instance, a rape, domestic violence, or child abuse case might elicit different results in which both intermediate and ultimate judgments are affected by stereotypic cues.

The likeability and knowledge manipulations have been well developed, used in prior research, and yielded successful manipulation checks in the present investigation. These successful experimental manipulations provide interpretive insight about the causal relations between expert gender, expert likeability, and expert knowledge on perceptions of credibility and case-related decision-making, results that cannot be obtained without sacrificing ecological validity to some extent. With regard to external validity, the

video-taped conditions were set in a realistic context, a witness stand in the Witness Research Lab, to provide the appearance of an authentic courtroom. Further, the expert witnesses were portrayed as forensic clinical mental health professionals with experience in testifying in court.

Although this study presents findings potentially relevant to men and women who testify in court, the study limitations may prevent the generalizability of the results to some extent. Our efforts to enhance external validity did not capture several important elements of capital trials. For instance, the dynamic of jury deliberation was not accounted for. Had the mock jurors discussed decision-making processes, the effect of our manipulations may have been different. Using a college student sample also may have limited the generalizability of the results. Studies that rely on undergraduate mock jurors to try to capture the trial process are limited by several internal and external validity concerns.⁴⁹ Consistent with the analysis by Wiener and colleagues,⁴⁹ these findings should be followed by an empirical examination of these hypotheses with a more representative sample and more realistic trial processes. Finally, although we conducted a manipulation check to see whether still images of the experts in this study elicited differential perceptions, future research should include multiple sample stimuli rather than a single instance of each stimulus category (e.g., at least two women and two men).⁵⁰

In conclusion, we found that likeability and knowledge are important for expert witness credibility, for both men and women. More research is needed to explore mock jurors' and actual jurors' cognitive responses in evaluating male and female expert witnesses. What is it that drives differential perceptions of male and female experts? Perhaps future research can provide more opportunities for both to demonstrate credible warmth and competence by allowing them to testify in cases with more stereotypical feminine and masculine domains.

Acknowledgments

Special thanks to Drs. Patricia Parmelee and James C. Hamilton for testifying as mock expert witnesses in this study and to undergraduate research assistants Matthew Jones, Katherine O'Brien, and Wyley Shreves for help with data collection and entry.

References

1. Eagly AH, Karau SJ: Role congruity theory of prejudice toward female leaders. *Psychol Rev* 109:573–98, 2002
2. Eagly AH, Makhijani MG, Klonsky BG: Gender and the evaluation of leaders: a meta-analysis. *Psychol Bull* 111:3–22, 1992
3. Guadagno RE, Cialdini RB: Gender differences in impression management in organizations: a qualitative review. *Sex Roles* 56: 483–94, 2007
4. Kipnis D, Schmidt S: Upward-influence styles: relationship with performance evaluations, salary, and stress. *Admin Sci Q* 33:528–42, 1988
5. Larson BA, Brodsky SL: When cross-examination offends: How men and women assess intrusive questioning of male and female expert witnesses. *J Appl Soc Psychol* 40:811–30, 2010
6. Neal TMS, Brodsky SL: Expert witness credibility as a function of eye contact behavior and gender. *Crim Just Behav* 35:1515–26, 2008
7. Schuller RA, Terry D, McKimmie B: The impact of expert testimony on jurors' decisions: gender of the expert and testimony complexity. *J Appl Soc Psychol* 35:1266–80, 2005
8. Fed. R. Evid. 702
9. Wojciszke B, Bazinska R, Jaworski M: On the dominance of moral categories in impression formation. *Pers Soc Psychol Bull* 24:1251–7, 1998
10. Fiske ST, Cuddy AJC, Glick P, *et al*: A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *J Pers Soc Psychol* 82: 878–902, 2002
11. Fiske ST, Xu J, Cuddy AJC, *et al*: (Dis)respecting versus (dis)liking: status and interdependence predict ambivalent stereotypes of competence and warmth. *J Soc Issues* 55:473–89, 1999
12. Cuddy AJC, Fiske ST, Kwan VSY, *et al*: Stereotype content model across cultures: towards universal similarities and some differences. *Br J Soc Psychol* 48:1–33, 2009
13. Brodsky SL, Griffin MP, Cramer RJ: The Witness Credibility Scale: an outcome measure for expert witness research. *Behav Sci Law* 28:892–907, 2010
14. Fiske ST, Cuddy AJC, Glick P: Universal dimensions of social cognition: warmth and competence. *Trends Cogn Sci* 11:77–83, 2006
15. Cuddy AJC, Fiske ST, Glick P: The BIAS map: behaviors from intergroup affect and stereotypes. *J Pers Soc Psychol* 9:631–48, 2007
16. Eagly AH: *Sex Differences in Social Behavior: A Social-Role Interpretation*. Hillsdale, NJ: Erlbaum, 1987
17. Cuddy AJC, Fiske ST, Glick P: When professionals become mothers, warmth doesn't cut the ice. *J Soc Issues* 60:701–18, 2004
18. Bridges J, Etaugh C, Barnes-Farrell J: Trait judgments of stay-at-home and employed parents: a function of social role and/or shifting standards? *Psychol Women Q* 26:140–50, 2002
19. Eagly AH, Steffen VJ: Gender stereotypes stem from the distribution of women and men into social roles. *J Pers Soc Psychol* 46:735–54, 1984
20. Eagly AH, Steffen VJ: Gender and aggressive behavior: a meta-analytic review of the social psychological literature. *Psychol Bull* 100:309–30, 1986
21. Eckes T: Paternalistic and envious gender stereotypes: testing predictions from the stereotype content model. *Sex Roles* 23:273–80, 2002
22. Etaugh C, Poertner P: Perceptions of women: influence on performance, marital and parental variables. *Sex Roles* 26:311–21, 1992
23. Etaugh C, Study GG: Perceptions of mothers: effects of employment status, marital status, and age of child. *Sex Roles* 20:59–70, 1989
24. Cramer RJ, Brodsky SL, DeCoster J: Expert witness confidence and juror personality: their impact on credibility and persuasion in the courtroom. *J Am Acad Psychiatry Law* 37:63–74, 2009

25. Brodsky SL, Neal TMS, Cramer RJ, et al: Credibility in the courtroom: how likeable should an expert witness be? *J Am Acad Psychiatry Law* 37:525–32, 2009
26. McAdams D, Powers J: Themes of intimacy in behavior and thought. *J Pers Soc Psychol* 40:573–87, 1981
27. Gladstone G, Parker G: When you're smiling does the whole world smile for you? *Australasian Psychiatry* 10:144–6, 2002
28. Kernis MH, Sun CR: Narcissism and reactions to interpersonal feedback. *J Res Pers* 28:4–13, 1994
29. Shuman HC: Attributes of gazing behaviour and physical contact in interpersonal relationships. *J Pers Clin Stud* 13:93–8, 1997
30. Levin H, Giles H, Garret P: The effects of lexical formality and accent on trait attributions. *Lang Comm* 14:265–74, 1994
31. Liska J, Hazelton V: Deferential language as a rhetorical strategy: the case for polite disagreement. *J Soc Behav Pers* 5:187–98, 1990
32. Kern J: Predicting the impact of assertive, empathic-assertive, and nonassertive behavior: the assertiveness of the assertee. *Behav Ther* 13:486–98, 1982
33. Champagne A, Shuman D, Whitaker E: An empirical examination of the use of expert witnesses in American courts. *Jurimetr J* 31:375–92, 1991
34. Ware J, Williams R: The Dr. Fox effect: a study of lecturer effectiveness and ratings of instruction. *J Med Educ* 50:149–56, 1975
35. Melton GB, Petrila J, Poythress NG, et al: *Psychological Evaluations for the Courts: A Handbook for Mental Health Professionals and Lawyers* (ed 3). New York: Guilford Press, 2007
36. Brodsky SL: *Testifying in Court: Guidelines and Maxims for the Expert Witness*. Washington, DC: American Psychological Association, 1991
37. Lee E: Categorical person perception in computer-mediated communication: effects of character representation and knowledge bias on sex inference and informational social influence. *Media Psychol* 9:309–29, 2007
38. *Witherspoon v. Illinois*, 391 U.S. 510 (1968)
39. Krauss DA, Sales BD: The effects of clinical and scientific expert testimony on juror decision making in capital sentencing. *Psychol Pub Policy Law* 7:267–310, 2001
40. Hosoda M, Stone-Romero EF, Coats G: The effects of physical attractiveness on job-related outcomes: a meta-analysis of experimental studies. *Personnel Psychol* 56:431–62, 2003
41. Ritts V, Patterson ML, Tubbs ME: Expectations, impressions, and judgments of physically attractive students: a review. *Rev Educ Res* 62:413–26, 1992
42. Schuller RA, Terry D, McKimmie B: The impact of an expert's gender on jurors' decisions. *Law Psychol Rev* 25:59–79, 2001
43. Adshad G: Evidence-based medicine and medicine-based evidence: the expert witness in cases of factitious disorder by proxy. *J Am Acad Psychiatry Law* 33:99–105, 2005
44. Chaiken S, Wood W, Eagly AH: Principles of persuasion, in *Social Psychology: Handbook of Basic Principles*. Edited by Higgins T, Kruglanski AW. New York: Guilford Press, 1996, pp 702–44
45. Petty RE, Cacioppo JT: The effects of involvement on responses to argument quantity and quality: central and peripheral approaches to persuasion. *J Pers Soc Psychol* 46:69–81, 1984
46. Greenwald AG: Cognitive learning, cognitive response to persuasion, and attitude change, in *Psychological Foundations of Attitudes*. Edited by Greenwald AG, Brock TC, Ostrom TM. New York: Academic Press, 1968, pp 147–70
47. Abshire J, Bornstein BH: Juror sensitivity to the cross-race effect. *Law Hum Behav* 27:471–80, 2003
48. Neal TMS, Christiansen A, Bornstein BH, et al: The effects of mock jurors' beliefs about eyewitness performance on trial judgments. *Psychol Crime Law* 18:49–64, 2012
49. Wiener RL, Krauss DA, Lieberman JD: Mock jury research: where do we go from here? *Behav Sci Law* 29:467–79, 2011
50. Wells GL, Windschitl PD: Stimulus sampling and social psychological experimentation. *Pers Soc Psychol Bull* 25:1115–25, 1999