

Neal, T.M.S., Guadagno, R.E., Eno, C.A., & Brodsky, S.L. (2012). Warmth and competence on the witness stand: Implications for credibility of male and female expert witnesses. *Journal of the American Academy of Psychiatry and the Law*, 40, 488-497.

This is a pre-print version of the paper that is available on the JAAPL website at <http://www.jaapl.org/content/40/4/488.full>. This article may not exactly replicate the final version published in the journal. It is not the copy of record.

© 2012 American Academy of Psychiatry and the Law

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

Tess M.S. Neal, M.A., Rosanna E. Guadagno, Ph.D.,  
Cassie A. Eno, Ph.D., & Stanley L. Brodsky, Ph.D.

Department of Psychology  
The University of Alabama

Tess M.S. Neal is now at Department of Psychiatry, University of Massachusetts Medical School. Cassie A. Eno is now affiliated with the Department of Psychology, Waldorf College. This research was presented at the 4<sup>th</sup> International Congress on Psychology and Law, Miami, Florida, 2011.

Correspondence concerning this article should be addressed to Tess M.S. Neal, University of Massachusetts Medical School, Department of Psychiatry WSH 8C, 55 Lake Avenue North, Worcester, MA 01655, USA. Phone: 508-856-5812. E-mail:

[tmneal@crimson.ua.edu](mailto:tmneal@crimson.ua.edu)

### Abstract

This study 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 credibility in the psycholegal domain. We found experts high in likeability and/or knowledge were perceived equally positively regardless of gender in a death penalty sentencing context. Gender differences emerged when the expert was low in likeability and/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.

*Keywords:* gender, credibility, witness, likeability, warmth, testimony

### Warmth and Competence on the Witness Stand:

#### Implications for Credibility of Male and Female Expert Witnesses

Women and men are perceived differently in many situations,<sup>1-4</sup> including when they serve as expert witnesses in legal settings.<sup>5-7</sup> 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.<sup>8</sup> 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 impacted by perceptions of expert witnesses. This study 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% of variance in global impressions.<sup>9</sup> 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.<sup>10-11</sup> Researchers from this perspective argue that we initially categorize individuals as being high or low in both of these domains. The stereotype content model<sup>10-11</sup> 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.<sup>12</sup>

The domains of warmth and credibility overlap conceptually with the *knowledge* and *likeability* subcomponents of witness credibility theory proposed by Brodsky, Griffin, and

Cramer.<sup>13</sup> Brodsky and colleagues 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).<sup>10-11</sup> People perceived as both warm and competent elicit uniformly positive emotions and behavior, while people perceived as lacking warmth and competence elicit uniform negativity.<sup>14</sup> People perceived as high on one dimension but low on the other elicit reliably ambivalent affective and behavioral responses.<sup>14</sup> For example, people perceived as high on competence but low on warmth elicit envy and competitive behaviors and people low on competence but high on warmth elicit pity and neglect.<sup>14-15</sup> 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.<sup>12</sup>

**Gender role expectations and perceptions.** According to social role theory, men and women experience different normative expectations for behavior.<sup>16</sup> Women are generally expected to be more warm and communal than men; that is, 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.<sup>3, 16-17</sup>

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<sup>18-22</sup> or as competent but cold.<sup>18, 21-23</sup> 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.<sup>14, 17</sup> Working women who become mothers do not gain perceived warmth after having children and in fact lose perceived competence.<sup>17</sup>

Although women perceived as high in both competence and warmth appear to be rare, Eno, Guadagno, and colleagues<sup>24</sup> 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 to being perceived 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 only be “successful” in two of the four combinations (e.g., cold/competent and warm/competent). That is, men who are perceived as incompetent may stand little chance to be perceived as credible, while women who are perceived as incompetent may still have a chance to be perceived 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 witnesses<sup>6</sup> 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 an assertive (high) eye contact level. These findings suggest men must be perceived as competent to enhance perceptions of credibility, but that women might be perceived 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 (e.g., warmth) and knowledge (e.g., competence) were manipulated. We use the terms “likeability” and “knowledge” because they have been established in the mental health-law literature;<sup>6, 13, 25-26</sup> however, the evidence reviewed above suggests that they overlap conceptually with “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.<sup>13, 26</sup> Thus, 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 behaviors consisted of the following components: consistent use of “we” or “us” when discussing members of the scientific community or humanity as a whole,<sup>27</sup> moderate levels of smiling,<sup>28</sup> modest statements and conclusions (e.g., “relatively certain” or “we do not know everything there is to know in psychology”),<sup>29</sup> consistent eye contact with lawyer and jury,<sup>30</sup> informal speech (i.e., low technical jargon and use of names),<sup>31</sup> and a self-effacing presentation style.<sup>32</sup>

**Low likeability:** Low likeable behaviors were 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 following the elements associated with high knowledge: degree of assertiveness,<sup>33</sup> substantive content and clarity of testimony,<sup>34-35</sup> credentials,<sup>8, 34, 36</sup> relevant experience,<sup>8, 36-37</sup> self-proclaimed expertise,<sup>38</sup> and familiarity with the case.<sup>34</sup> We also manipulated conceptions of high and low knowledge.

**High knowledge:** High levels of knowledge were demonstrated by strong educational credentials (e.g., educated at Yale, ABPP certified in Forensic Psychology, history of academic publication in this area of expertise),<sup>8, 34, 36</sup> solid relevant experience (e.g., risk assessment researcher, conducted over 100 such assessments over 14 years)<sup>8, 36-37</sup> consistent clarity and substantive content of communication,<sup>34-35</sup> moderate assertiveness (e.g., “as far as I know I've never been wrong” when queried about awareness of clinician error),<sup>33</sup> self-proclaimed expertise (e.g., “In my expert opinion...”),<sup>38</sup> and demonstrated familiarity with the case (e.g., multiple interviews with the defendant).<sup>34</sup>

**Low Knowledge:** The variable of low knowledge consisted of the following behaviors: no mention of educational credentials, minimal relevant experience (e.g., 2 years as a



psychotherapist, no previous experience in risk assessment), inconsistent clarity and substantive content of communication, low 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:

1. The most credible experts and those eliciting the highest agreement ratings from mock jurors would be experts high in both likeability and knowledge.
2. The least credible experts with the lowest agreement ratings would be those low in both likeability and knowledge.
3. In the high likeability condition female experts would be perceived as more credible and have higher agreement ratings than male experts, but in the low likeability condition male experts would be perceived as more credible with higher agreement ratings than female experts.
4. No interaction between gender and knowledge of expert.
5. Female experts high in likeability but low in knowledge would be perceived as equally credible and elicit equal agreement ratings compared to male experts.  
Although SCM would predict women would be more credible than men with this combination of likeability and knowledge,<sup>17</sup> social role theory would predict the man would be perceived as more competent, due to the role of an expert as a masculine role. Therefore, we predicted both theories would be true, washing out any difference for this particular combination.
6. Males and females high in knowledge but low in likeability would be perceived as equally credible and would elicit equal agreement ratings.

## Method

### Participants

Undergraduate students at a large public university in Alabama, U.S.A. were recruited through the Psychology Subject Pool and participated for course credit ( $N=265$ ). The gender composition of our sample was evenly split (51% female) and ranged in age from 18 to 36 years ( $M=19.64$ ,  $SD=2.01$ ). Seventy-six percent were Caucasian, 16% were African-American, and 8% from other ethnic backgrounds. The Supreme Court decided in *Witherspoon vs. Illinois*<sup>39</sup> that jurors who sit on capital murder trials must be “death qualified;” that is, they must be willing and able to consider capital punishment as an appropriate 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 the eligible participants on any demographic variable in our dataset. It should be noted that Alabama, the state in which 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) x 2 (high vs. low likeability) x 2 (high vs. low knowledge) between-subjects factorial design. We developed eight videos to match the eight conditions of the study, each of which was approximately 7 minutes long. The script for the videos was adapted from an actual jury sentencing proceeding described by Krauss and Sales.<sup>40</sup> Previous research has successfully used the same basic script to examine expert witness credibility.<sup>6, 25-26</sup> 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.<sup>13</sup> The scale contains 20 bi-polar adjectives on a 10-point Likert scale. Higher scores indicate greater credibility ratings. A few examples of these bi-polar adjectives include: “*unkind*” (1) to “*kind*” (10); “*dishonest*” (1) to “*honest*” (10); and “*shaken*” (1) to “*poised*” (10). Factor analysis identified four independent domains of trustworthiness, confidence, likeability, and knowledge.<sup>13</sup> 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 of the defendant re-offending, 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 ten-point Likert-type scales: (a) How likeable did you find this expert witness? (“*not at all likeable*” [1] to “*extremely likeable*” [10]), (b) How knowledgeable did you find this expert witness? (“*not at all knowledgeable*” [1] to “*extremely knowledgeable*” [10]) and c) How physically attractive did you find this expert witness (“*not at all attractive*” 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.<sup>41-42</sup>

## **Procedure**

As per the approved Institutional Review Board protocol for this study, 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,<sup>13</sup> a sentencing rating form which 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, participants were debriefed.

## **Results**

### **Pilot Study**

We conducted a between-subjects MANOVA to compare mock juror ratings ( $n=22$ ) of still photographic images of each expert witnesses to ensure the experts were matched on

credibility, likeability, knowledge, attractiveness, jurors' substantive agreement with them, and recommended juror sentence prior to commencing data collection with the videos. Results indicated there were no significant differences based on the images of the male and female experts,  $Wilks' \Lambda = 0.73$ ,  $F(8, 14) = 0.65$ ,  $p = 0.73$ ,  $\eta_p^2 = 0.27$ .

### Primary Analyses

**Manipulation Check.** Our manipulation check items indicated our manipulations of knowledge and likeability were successful for each expert; that is, the knowledge manipulation worked,  $F(1, 262) = 25.79$ ,  $p < 0.001$  (high knowledge  $M = 86.92$ ,  $SD = 14.58$  vs. low knowledge  $M = 75.26$ ,  $SD = 21.40$ ), as did the likeability manipulation,  $F(1, 262) = 226.76$ ,  $p < 0.001$  (high likeability  $M = 72.24$ ,  $SD = 25.49$  vs. low likeability  $M = 23.34$ ,  $SD = 27.18$ ). Further support indicating we successfully manipulated these individual constructs is that the likeability manipulation did not affect knowledge ratings,  $F(1, 262) = 2.19$ ,  $p = 0.14$  and the knowledge manipulation did not affect likeability ratings,  $F(1, 262) = 0.03$ ,  $p = 0.87$ .

**Main Analyses.** To conduct our primary analyses, we conducted a MANOVA with our three independent variables (gender: male vs. female expert; likeability condition: high vs. low likeability; 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, so we did not include them in our final models. Significant multivariate main effects emerged for knowledge condition,  $Wilks' \Lambda = 0.95$ ,  $F(3, 231) = 4.05$ ,  $p = 0.008$ ,  $\eta_p^2 = 0.05$ , and likeability condition,  $Wilks' \Lambda = 0.71$ ,  $F(3, 231) = 31.71$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.29$ . 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 below. The main effect of gender of expert was not significant, *Wilks' Lambda* = 0.97,  $F(3, 231) = 2.07$ ,  $p = 0.11$ ,  $\eta_p^2 = 0.03$ , indicating 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,  $F(1, 233) = 3.69$ ,  $p = 0.05$ ,  $\eta_p^2 = 0.02$ . 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 three other combinations of likeability and knowledge ( $M$  range = 6.22 – 7.71,  $SD$  range = 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 likeability or knowledge of the experts.

Our second hypothesis, that the least credible experts would be those low in both likeability and knowledge, was also supported,  $F(1, 233) = 9.29$ ,  $p = 0.003$ ,  $\eta_p^2 = 0.04$ . 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 three other combinations of likeability and knowledge ( $M$  range = 6.81 – 8.14,  $SD$  range = 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 the likeability of the expert would differentially affect credibility ratings for 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 ( $M = 6.69$ ,  $SD = 1.18$ ),  $F(1, 233) = 4.96$ ,  $p = 0.027$ ,  $\eta_p^2 = 0.02$ , and participants agreed more with the male expert ( $M = 73.37$ ,  $SD = 17.50$ )

than with the female expert ( $M = 66.31, SD = 15.43$ ),  $F(1, 233) = 5.92, p = 0.016, \eta_p^2 = 0.03$ .

Although we expected that the female expert would be perceived 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,  $F(1, 233) = 2.93, p = 0.09, \eta_p^2 = 0.01$ .

Hypothesis four predicted no interaction between gender of expert and knowledge. According to the MANOVA results described above, this hypothesis was supported: no significant effects were found for our collection of dependent variables, *Wilks' Lambda* = 0.99,  $F(3, 231) = 0.74, p = 0.53, \eta_p^2 = 0.01$ . However, we noted an interesting pattern to the univariate results, which we describe in the exploratory findings section below.

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 ( $M = 7.52, SD = 1.32$ ),  $F(1, 233) = 1.68, p = 0.20, \eta_p^2 = 0.07$ , and no differences emerged in percentage of agreement with the male expert ( $M = 76.12, SD = 16.37$ ) and female expert ( $M = 70.34, SD = 17.34$ ),  $F(1, 233) = 1.47, p = 0.23, \eta_p^2 = 0.01$ .

We predicted that highly knowledgeable but unlikeable experts, whether they were men or women, would be perceived as equally credible and would elicit similar ratings of substantive agreement. The data support this hypothesis for credibility ratings,  $F(1, 233) = 1.01, p = 0.32, \eta_p^2 = 0.01$ , 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,  $F(1, 233) = 2.04, p = 0.16, \eta_p^2 = 0.01$ , with no differences in agreement with the male

( $M = 73.79$ ,  $SD = 15.39$ ) versus female ( $M = 67.68$ ,  $SD = 15.37$ ) experts. An interesting note is that the woman was never rated as more credible than the male 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 to conduct 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 expert would differ in the low knowledge and low 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 Figure 1). For instance, participants rated the low knowledgeable male expert as significantly more credible ( $M = 7.06$ ,  $SD = 1.41$ ) than the comparable female ( $M = 6.71$ ,  $SD = 1.53$ ),  $F(1, 233) = 5.94$ ,  $p = 0.016$ ,  $\eta_p^2 = 0.03$  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 ( $M = 68.02$ ,  $SD = 16.66$ ),  $F(1, 233) = 5.16$ ,  $p = 0.024$ ,  $\eta_p^2 = 0.02$ . 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 ( $M = 5.87$ ,  $SD = 1.25$ ),  $F(1, 233) = 4.83$ ,  $p = 0.029$ ,  $\eta_p^2 = 0.02$ .



Likewise, participants agreed significantly more with the substantive content of the low likeable and low in knowledge male expert ( $M = 73.05$ ,  $SD = 19.11$ ) compared to the corresponding female expert ( $M = 65.03$ ,  $SD = 15.65$ ),  $F(1, 233) = 4.17$ ,  $p = 0.042$ ,  $\eta_p^2 = 0.02$  (see Figure 1).

[Insert Figure 1 about here]

While conducting our data analyses, we noticed that 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.

Results indicated that the full  $R^2 = 0.602$ , indicating that the model (i.e., regressing the likeability and knowledge ratings on total credibility) predicted 60.2% of the variability in total credibility score. Thus, likeability and knowledge jointly explained a significant amount of variability in the total credibility score,  $F(2, 238) = 180.30$ ,  $p < 0.001$ . 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,  $F(1, 238) = 55.90$ ,  $p < 0.001$ . 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.<sup>9-13</sup> Consistent with this literature, we found that experts high in likeability and knowledge were perceived as the most credible, while 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 psycholegal domain.

With respect to gender differences in evaluations of our expert witnesses, our results revealed that participants evaluated men and women differently when a) in the role of an expert; and b) when appearing low in likeability and/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 poorly compared to men, consistent with prior literature. We take a social role theory perspective<sup>16</sup> on these findings and suggest that this was the case because the woman was in a masculine occupational role – 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 “feminine” (e.g., perhaps child abuse), the pattern of results may have differed.<sup>43, 44</sup>

This rationale is further supported by the fact that our pilot study found that without the occupational role, participants evaluated the experts’ similarly. The fact that we included only one domain of testimony – violent recidivism in a capital murder sentencing context – is a

limitation of this study. The SCM<sup>14, 17</sup> predicts that women need to be perceived as both warm and competent to be compared favorably to men. Perhaps future research can provide more opportunities for men and women experts to demonstrate credible warmth and competence by allowing the experts to testify in cases with more stereotypical feminine as well as masculine domains.

The range of credibility assessments for expert knowledge in this study was relatively restricted, which suggests participants accepted the witness as an “expert” and perceived him/her as knowledgeable even when his/her 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.<sup>45-46</sup> 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 responses<sup>47</sup> 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 expert.<sup>46</sup> Alternatively, experts low in knowledge and likeability may have still reached the minimum threshold for mock jurors to evaluate them as experts. This pattern may have particularly been likely given that the expert witnesses were in their mid-50’s 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 expected results (i.e., credibility ratings and substantive agreement), when it came to the ultimate decision a juror must make (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.<sup>48-49</sup> 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 (discussed in more detail below). 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 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. The efforts to enhance external validity described above 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 these results. Studies that rely on undergraduate mock jurors to try to capture the trial process are limited by several internal and external validity concerns.<sup>50</sup> Consistent with Wiener and colleagues' analysis, 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).<sup>51</sup>

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 men and women experts? Perhaps future research can provide more opportunities for men and women experts to demonstrate credible warmth and competence by allowing the experts to testify in cases with more stereotypical feminine as well as masculine domains.

### Acknowledgement

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 their help with data collection and entry.

## References

1. Eagly AH, Karau S J: Role congruity theory of prejudice toward female leaders. *Psychol Rev*, 109: 573-98, 2002
2. Eagly AH, Makhijani M G, 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 Applied 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 and 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 Applied 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. *Personality and 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 Personality and 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. *Brit J of 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 & the L*, 28: 892-907, 2010
14. Fiske ST, Cuddy AJC, Glick P: Universal dimensions of social cognition: Warmth and competence. *TRENDS in Cognitive Sci*, 11:, 77-83, 2006
15. Cuddy AJC, Fiske ST, Glick P: The BIAS Map: Behaviors from intergroup affect and stereotypes. *J of Personality and 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 of 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 of 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 of Personality and 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. Eno CA, Guadagno RE, Habashi MH, *et al*: The hockey mom: A new subtype for warm and competent women. Manuscript Under Review, 2012
25. Cramer, RJ, Brodsky, SL, DeCoster, J: Expert witness confidence and juror personality: Their impact on credibility and persuasion in the courtroom. *J of the Am Acad of Psychiatry and the L*, 37: 63-74, 2009
26. Brodsky SL, Neal TMS, Cramer, RJ, *et al*: Credibility in the courtroom: How likeable should an expert witness be? *J of the Am Acad of Psychiatry and the L*, 37: 525-32, 2009
27. McAdams D, Powers J: Themes of intimacy in behavior and thought. *J Personality and Soc Psychol*, 40: 573-87, 1981
28. Gladstone G, Parker G: When you're smiling does the whole world smile for you? *Australasian Psychiatry*, 10: 144-6, 2002
29. Kernis MH, Sun CR: Narcissism and reactions to interpersonal feedback. *J Res in Personality*, 28: 4-13, 1994
30. Shuman HC: Attributes of gazing behaviour and physical contact in interpersonal relationships. *J Personality and Clinical Studies*, 13: 93-8, 1997
31. Levin H, Giles H, Garret P: The effects of lexical formality and accent on trait attributions. *Language & Comm*, 14: 265-74, 1994.

32. Liska J, Hazelton V: Deferential language as a rhetorical strategy: The case for polite disagreement. *J Soc Behav & Personality*, 5: 187-98, 1990
33. Kern J: Predicting the impact of assertive, empathic-assertive, and nonassertive behavior: The assertiveness of the assertee. *Behav Therapy*, 13: 486-98, 1982
34. Champagne A, Shuman D, Whitaker E: An empirical examination of the use of expert witnesses in American courts. *Jurimetrics J*, 31: 375-92, 1991
35. Ware J, Williams R: The Dr. Fox effect: A study of lecturer effectiveness and ratings of instruction. *J Medical Education*, 50: 149-56, 1975
36. Melton GB, Petrila J, Poythress NG, *et al*: *Psychological Evaluations for the Courts: A Handbook for Mental Health Professionals and Lawyers*, Third Edition. New York: Guilford Press, 2007
37. Brodsky SL: *Testifying in Court: Guidelines and Maxims for the Expert Witness*. Washington DC: American Psychological Association, 1991
38. 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
39. *Witherspoon v. Illinois*, 391 U.S. 510 (1968)
40. Krauss DA, Sales BD: The effects of clinical and scientific expert testimony on juror decision making in capital sentencing. *Psychol, Pub Pol'y, & L*, 7: 267-310, 2001
41. 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
42. Ritts V, Patterson ML, Tubbs ME: Expectations, impressions, and judgments of physically attractive students: A review. *Rev Educational Res*, 62: 413-26, 1992

43. Schuller RA, Terry D, McKimmie B: The impact of an expert's gender on jurors' decisions. *Law and Psychol Rev*, 25: 59-79, 2001
44. Adshead G: Evidence-based medicine and medicine-based evidence: The expert witness in cases of factitious disorder by proxy. *J of the Am Acad of Psychiatry and the L*, 33: 99-105, 2005
45. 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
46. Petty RE, Cacioppo JT: The effects of involvement on responses to argument quantity and quality: Central and peripheral approaches to persuasion. *J Personality and Soc Psychol*, 46: 69-81, 1984
47. 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
48. Abshire J, Bornstein BH: Juror sensitivity to the cross-race effect. *L and Hum Behav* 27: 471-80, 2003
49. Neal TMS, Christiansen A, Bornstein BH, *et al*: The effects of mock jurors' beliefs about eyewitness performance on trial judgments. *Psychol, Crime, & L*, 18: 49-64, 2012
50. Wiener RL, Krauss DA, Lieberman JD: Mock jury research: Where do we go from here? *Behav Sci & the L*, 29: 467-79, 2011
51. Wells GL, Windschitl PD: Stimulus sampling and social psychological experimentation. *Personality and Soc Psychol Bull*, 25: 1115-25, 1999

Figure 1. Interaction between expert gender, knowledge, and likeability on credibility ratings and percent agreement with expert.

