Explaining Gender Differences in Jurors’ Reactions to Child Sexual Assault Cases

Bette L. Bottoms*, †, Liana C. Peter-Hagene‡, Margaret C. Stevenson†,‡, Tisha R. A. Wiley†, Tracey Schneider Mitchell§ and Gail S. Goodman¶

In three experiments, we investigated the influence of juror, victim, and case factors on mock jurors’ decisions in several types of child sexual assault cases (incest, day care, stranger abduction, and teacher-perpetrated abuse). We also validated and tested the ability of several scales measuring empathy for child victims, children’s believability, and opposition to adult/child sex, to mediate the effect of jurors’ gender on case judgments. Supporting a theoretical model derived from research on the perceived credibility of adult rape victims, women compared to men were more empathic toward child victims, more opposed to adult/child sex, more pro-women, and more inclined to believe children generally. In turn, women (versus men) made more pro-victim judgments in hypothetical abuse cases; that is, attitudes and empathy generally mediated this juror gender effect that is pervasive in this literature. The experiments also revealed that strength of case evidence is a powerful factor in determining judgments, and that teen victims (14 years old) are blamed more for sexual abuse than are younger children (5 years old), but that perceptions of 5 and 10 year olds are largely similar. Our last experiment illustrated that our findings of mediation generalize to a community member sample. Copyright © 2014 John Wiley & Sons, Ltd.

Many children testify in the courts of law, especially in child sexual assault cases (Bulkley, 1989; Goodman, Quas, Bulkley, & Shapiro, 1999). Given the inherently secretive nature of child sexual abuse, jurors often struggle to make decisions based on little evidence other than the word of a child (Myers, 1998). In fact, jurors in child maltreatment trials report that they consider child testimony the most important evidence (Myers et al., 1999). In such cases with little “hard” evidence, extralegal factors can have a particularly strong influence.

A robust finding in the literature on jurors’ decisions in child sexual abuse trials is that women assign more blame and are generally more prosecution-oriented than men (e.g., Bottoms & Goodman, 1994; Kovera, Levy, Borgida, & Penrod, 1994; McCauley & Parker, 2001; Quas, Bottoms, Haegerich, & Nysses-Carras, 2002; Redlich, Myers, Goodman, & Qin, 2002; for a comprehensive review, see Bottoms et al., 2007). In three experiments, we investigated factors explaining those gender differences across four types of cases (incest, day-care, stranger abduction, and

---

*Correspondence to: Bette L. Bottoms, or by mail at University of Illinois at Chicago, Department of Psychology (MC 285), 1007 W. Harrison St., Chicago, IL 60607-7137. E-mail: bbottoms@uic.edu
†University of Illinois at Chicago, IL
‡University of Evansville, IN
§State University of New York at Buffalo, NY
¶University of California, Davis, CA

Copyright © 2014 John Wiley & Sons, Ltd.
teacher-perpetrated sexual abuse). We theorized that gender differences would be explained by underlying differences in attitudes and empathy toward child victims. To this purpose, we developed and validated three scales: Children’s Believability Scale, Opposition to Adult/Child Sex, and Child Victim Empathy. We then tested our main hypotheses that these scales would: (a) predict case judgments; and (b) mediate the effects of juror gender on case judgments. In addition, we investigated the effects of several other important factors on case judgments: child victim age (Experiments 1 and 2), case strength (Experiment 2), and sample type – community member versus undergraduate (Experiment 3).

**Attitudinal Predictors**

**Attitudes Toward Children’s Believability**

Extrapolation from various literatures suggests several attitudinal constructs that might influence decision-making in child sexual assault cases. For example, beliefs about suggestibility and cognitive deficits lead to decreased perceived credibility for child bystander witnesses (Goodman, Golding, & Haith, 1984; Leippe & Romanczyk, 1987, 1989), but to increased credibility in cases of sexual victimization, in which: (a) a child’s sincerity and trustworthiness are highlighted more than cognitive abilities; and (b) a perceived lack of cognitive sophistication is associated with an inability to fabricate sexual allegations (e.g., Bottoms & Goodman, 1994; Goodman et al., 1984). Gabora, Spanos, and Joab (1991, 1993) found that mock jurors who endorsed children’s proneness to fabricate false allegations of sexual assault (on the child sexual abuse belief scale) were less likely than others to believe a child and less likely to vote guilty.

Further, there are gender differences in such attitudes: Bull et al. (1991) reported that, compared with men, women were more likely to think that children have reasonable witness abilities; can accurately identify their assailants; have memories for trauma that rival that of adults; and are not prone to sexual fantasy. Bottoms and Goodman (1994) uncovered that, compared with men, women considered alleged child sexual assault victims to be less suggestible. Gabora et al. (1993) found that men were more likely than women to endorse children’s proneness to fabricate false allegations of sexual assault.

In summary, attitudes toward child sexual assault victims’ believability consist of beliefs about children’s cognitive abilities and children’s tendencies to fabricate allegations, there are gender differences in these attitudes, and these attitudes are linked to child sexual abuse case judgments. Therefore, we predicted that gender differences in attitudes toward children’s believability would contribute to gender differences in child sexual assault case judgments.

**Opposition to Adult/Child Sex**

Collings (1997) documented child sexual abuse myths in a similar manner to adult rape myths (i.e., victim blame, denial of the crime and its seriousness). Just as attitudes toward rape are related to evaluations made in adult rape cases (e.g., Burt & Albin, 1981; Feild, 1978), jurors’ attitudes toward child sexual abuse might predict their case judgments, regardless of laws mandating that, for example, issues of consent must be ignored in child sexual abuse cases. Further, as is true for adult sexual abuse (e.g., White & Kurpius, 1999), men and women differ in their
attitudes toward child sexual abuse (e.g., Bull et al., 1991; Finlayson & Koocher, 1991; Johnson, Owens, Dewey, & Eisenberg, 1990), with women (a) viewing child sexual abuse as more harmful and more widespread, (b) being less likely than men to attribute abuse to victim’s sexual behavior, and (c) being less likely to believe children are prone to sexual fantasy. Therefore, we predicted that gender differences in attitudes toward child sexual abuse would explain gender differences in reactions to specific child sexual abuse cases.

**Empathy**

Empathy is the ability to understand another’s perspective, cognitively and emotionally (Davis, 1980, 1983). Compared with men, women are generally more empathic (Barnett et al., 1992; Batson et al., 1988; Davis, 1983; but see Eisenberg, Fabes, Schaller, & Miller, 1989). Deitz, Blackwell, Daley, and Bentley (1982) found that jurors with more (vs. less) rape victim empathy make more pro-victim judgments in adult rape cases, and further, that men have less rape victim empathy than do women.

Thus, we expected child sexual assault case judgments to be related to jurors’ empathy for child victims. We defined such empathy as “the tendency to identify cognitively and emotionally with child victims of sexual assault, having feelings of concern, compassion, and understanding for their experience,” a definition informed by Batson, Fultz, and Schoenrade’s (1987) work. We expected jurors with more (vs. less) child victim empathy to make pro-prosecution case decisions, because emotional and cognitive identification with child victims would predispose them to favor child victims (perhaps protectively), to attribute less blame to children, and to want to punish their abusers. Further, we expected that gender differences in child victim empathy would underlie gender differences in case judgments. In support, Leippe, Brigham, Cousins, and Romanczyk (1989) found that, compared with men, female attorneys thought it more likely that children make retractions of sexual abuse allegations because of embarrassment or fear of retaliation rather than because initial reports are false. Leippe et al. (1989) interpreted this as evidence for women’s more empathic approach to children (see also Bederian-Gardner & Goldfarb, 2014, pp. 829–845).

**Overview**

The major goal of all three experiments was to develop reliable scale measures of the proposed constructs (children’s believability, opposition to adult/child sex, and child victim empathy) and to test their predictive validity across scenarios representing different types of child sexual assault cases. In each experiment, participants completed scale measures and assumed the role of juror, reading case summaries and making case judgments, including verdict. We predicted that our measures would be reliable and that attitudes and empathy would mediate the effect of jurors’ gender on case judgments as summarized earlier. We also varied and measured other variables in the three experiments. Next, we describe the rationale for each as we detail the experiments, their methods, and results.
EXPERIMENT 1

In Experiment 1, we constructed and assessed the reliability and discriminant validity of our three constructed scales, and then tested our hypotheses about their predictive validity in a mock trial study. We expected that effects would generalize across different case types and cases involving children of different ages. Participants completed the Interpersonal Reactivity Index (IRI; Davis, 1980, 1983), a self-report measure of trait empathy, to determine the discriminant validity of the Child Victim Empathy Scale and to test the ability of a more general trait empathy measure to predict judgments as compared with the more specific Child Victim Empathy Scale.1

We also tested the influence of child victim age on case judgments. Older children are generally perceived as less credible than younger children in child sexual abuse cases (e.g., Bottoms, Davis, & Epstein, 2004; Golding, Sanchez, & Sego, 1999; but see Golding, Fryman, Marsil, & Yozwiak, 2003; McCauley & Parker, 2001; for a review, see Bottoms et al., 2007). Bottoms and Goodman (1994) theorized that, compared with older children and adults, young children are generally perceived to be low in competence (cognitive ability, resistance to suggestion), yet high in trustworthiness (honesty, sincerity, innocence; see also Goodman et al., 1984; Leippe & Romanczyk, 1987; Leippe et al., 1989), and that jurors attribute credibility to young child sexual abuse victims because they view these victims as honest and trustworthy, but also as sexually naive, lacking the knowledge and cognitive capacity to fabricate sexual encounters. In addition, young children are considered unlikely to possess characteristics that lower adult rape victims’ credibility, such as previous sexual experience, sexual provocativeness, and ability to give meaningful consent (e.g., Borgida & Brekke, 1985).

For these reasons, mock trial studies show that very young children are often considered more credible than children over the age of about 13 years old. In Experiment 1, we varied child age from 5 to 10 years, assuming that the 5-year-old would be perceived as particularly credible because of presumed honesty and naivety, but not as cognitively competent as the 10-year-old, who would probably also still be young enough to be perceived as naive and honest. Thus, we predicted that jurors would be less likely to convict defendants in cases involving the younger than the older (5 vs. 10 years old) victim.

Method

Experiment 1 conformed to a 2 (juror gender: men, women) × 2 (Victim age: 5, 10 years) × 4 (case type: incest, day-care abuse, stranger abduction, teacher assault) factorial design, with victim age varied between subjects and case type varied within subjects.

1 The Juror Bias Scale (Kassin & Wrightsman, 1996), Sociosexual Orientation Inventory (Simpson & Gangestad, 1991), and Marlowe-Crowne Social Desirability Scale (MCSDS; Crowne & Marlowe, 1964) were also included to demonstrate discriminant validity from general juror biases, general sexual conservatism, and social desirability response bias, respectively. Correlations between these and our created scales were not high. Further, the effects of our new scales remained significant in additional analyses controlling for juror bias, sexual conservatism, and desirability.
Participants

Participants were 132 introductory psychology students at a large northeastern public research institution (48% women; $M_{age} = 19$ years; 83% White, 8.5% African American, 4.5% Asian American, 4% “other”). All were United States citizens.

Materials

Demographic Questions. Questions assessed participant age, race/ethnicity, and gender.

Attitudes Questionnaire (APPENDIX). The 46-item attitude questionnaire was composed of randomly ordered items designed to measure the three proposed constructs: (a) children’s believability (e.g., “Children are not capable of inventing stories of sexual abuse”); (b) opposition to adult/child sex (e.g., “In some cases, children invite sexual contact with adults,” reverse-coded); and (c) child victim empathy (e.g., “It makes me sad to hear about children who have been sexually molested”). All items were measured with seven-point Likert scales ranging from −3 (strongly disagree) to +3 (strongly agree), and approximately half were reverse-scored to avoid response bias. Scores were recoded for analyses (1 = strongly disagree to 7 = strongly agree). Instructions specified that items mentioning a child or children referred to 5- to 10-year-olds. Items were derived from a review of related empirical and theoretical literature and constructed for this study, with the exception of six empathy items and three sexuality items modeled after items from the Rape Empathy Scale (RES; Deitz et al., 1982).

IRI. This multidimensional empathy measure (Davis, 1980, 1983) consists of four seven-item subscales designed to tap several aspects of empathy: perspective-taking, fantasy, empathic concern, and personal distress (e.g., “I often have tender, concerned feelings for people less fortunate than me”). The five-point response scale ranged from 0 (does not describe me well) to 4 (describes me very well). Davis (1980) reports satisfactory internal (0.71–0.77) and test–retest (0.62–0.71) reliabilities; the scale was also reliable in our sample ($\alpha = 0.84$).

Case Scenarios and Judgment Scales. Four one-page scenarios described a 5- or 10-year-old (all the same age) girl victim of four different types of child sexual assault situations that were modeled on actual cases of incest, teacher–student abuse, day-care abuse, and stranger abduction (Bottoms, 1992). Scenario order was counterbalanced according to a Latin square design. For each scenario, the judgment scales started with a dichotomous rating of the guilt/innocence of the defendant (guilty/not guilty), followed by confidence in the verdict three-point scale ranging from 1 (not at all confident) to 3 (very confident). The combination of these ratings resulted in a six-point “degree of guilt” scale ranging from 1 (not guilty/very confident) to 6 (guilty/very confident) ($M = 4.01$, $SD = 1.67$). The degree-of-guilt scale is a commonly used supplemental measure of jurors’ verdict tendencies (e.g., Kassin, Rigby, & Castillo, 1991; Krauss, McCabe, & Lieberman, 2012) that is more sensitive than a dichotomous verdict.

Procedure

Students participated in two experimental sessions. Session 1 was a general mass testing session during which participants completed the attitude questionnaire and the IRI, as well as other unrelated scale measures. Two months later, a random sample of those
students participated in the ostensibly unrelated Session 2: In groups of 10–40, participants were asked to play the role of mock jurors and read each of the four case scenarios (with child age varied between subjects). Participants were instructed about the importance of the research and asked to be thoughtful and serious in completing the packet; they then made judgments concerning the cases, and completed all other measures, including the attitude questionnaire for a second time. Participants were thanked and debriefed.

Results

Scale Construction and Psychometric Properties

Reliability. Separate Cronbach α analyses were conducted on the multi-item scales for Session 1 and Session 2 (see Table 1 for means and SDs). Several items were dropped from their respective scales because their inclusion weakened scale reliability (α and mean inter-item correlations) and because their item-total correlations were low. Inter-scale correlations were significant, yet not as high as individual scale reliabilities, suggesting that the scales tapped distinct constructs (see Table 2). Remaining items had acceptable item-total correlations and formed scales with acceptable internal consistencies (α-values =0.71–0.86; mean inter-item correlations =0.19–0.34) at Session 1 and Session 2 (see APPENDIX). Test–retest reliability, determined by correlating scale scores from Sessions 1 and 2, was also acceptable: for Child Victim Empathy, \( r = 0.69 \); for Opposition to Adult/Child Sex, \( r = 0.67 \); for Children’s General Believability, \( r = 0.66 \).

Discriminant Validity. There was a moderate, significant correlation between the IRI and the Child Victim Empathy (\( r = 0.32 \)) and Children Believability scales.

<table>
<thead>
<tr>
<th>Scales</th>
<th>Overall</th>
<th>Men</th>
<th>Women</th>
<th>t-test values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children’s Believability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session 1</td>
<td>4.39 (0.81)</td>
<td>4.13 (0.76)</td>
<td>4.66 (0.78)</td>
<td>( t(130) = 4.01 )</td>
</tr>
<tr>
<td>Session 2</td>
<td>4.05 (0.94)</td>
<td>3.68 (0.87)</td>
<td>4.44 (0.87)</td>
<td>( t(130) = 4.98 )</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>4.11 (1.01)</td>
<td>3.80 (0.84)</td>
<td>4.43 (1.07)</td>
<td>( t(390) = 6.48 )</td>
</tr>
<tr>
<td>Experiment 3</td>
<td>3.87 (1.08)</td>
<td>3.66 (1.06)</td>
<td>4.03 (1.07)</td>
<td>( t(437) = 3.77 )</td>
</tr>
<tr>
<td>Child Victim Empathy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session 1</td>
<td>5.91 (0.62)</td>
<td>5.65 (0.63)</td>
<td>6.19 (0.64)</td>
<td>( t(130) = 5.46 )</td>
</tr>
<tr>
<td>Session 2</td>
<td>5.92 (0.63)</td>
<td>5.61 (0.60)</td>
<td>6.25 (0.48)</td>
<td>( t(130) = 6.70 )</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>5.88 (0.64)</td>
<td>5.63 (0.65)</td>
<td>6.14 (0.51)</td>
<td>( t(390) = 8.65 )</td>
</tr>
<tr>
<td>Experiment 3</td>
<td>6.02 (0.69)</td>
<td>5.77 (0.73)</td>
<td>6.20 (0.61)</td>
<td>( t(387) = 6.94 )</td>
</tr>
<tr>
<td>Opposition to Adult/Child Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session 1</td>
<td>5.53 (0.76)</td>
<td>5.25 (0.79)</td>
<td>5.84 (0.59)</td>
<td>( t(130) = 4.88 )</td>
</tr>
<tr>
<td>Session 2</td>
<td>5.62 (0.79)</td>
<td>5.32 (0.81)</td>
<td>5.94 (0.65)</td>
<td>( t(130) = 4.83 )</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>5.48 (0.86)</td>
<td>5.29 (0.89)</td>
<td>5.67 (0.79)</td>
<td>( t(390) = 4.46 )</td>
</tr>
<tr>
<td>Experiment 3</td>
<td>5.43 (0.93)</td>
<td>5.25 (0.97)</td>
<td>5.56 (0.88)</td>
<td>( t(410) = 3.53 )</td>
</tr>
<tr>
<td>Attitudes Toward Women</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2</td>
<td>3.29 (0.43)</td>
<td>3.06 (0.44)</td>
<td>3.53 (0.25)</td>
<td>( t(389) = 12.85 )</td>
</tr>
</tbody>
</table>

Note. All significant at \( p < 0.01 \).
(r =0.39); however, the relations are not so high as to indicate that the Child Victim Empathy Scale measures the same construct of general empathy purportedly measured by the IRI.

**Scale Scores, Gender, and Case Judgments.** Compared with men, women were significantly more empathic toward child victims, more believing in them, and more opposed to adult/child sex at both sessions (see Table 2 for all means). Overall, verdicts were well balanced between not guilty (40%) and guilty (60%), with significantly fewer men (54%) than women (68%) voting guilty [χ² (N=525) =8.68, p =0.003]. Similarly, an analysis of variance (ANOVA) varying juror gender, victim age (between), and case type (within) revealed that women (M=4.30, SD =1.59) endorsed significantly higher degree-of-guilt than did men (M =3.74, SD =1.69) [F(1, 126) =12.01, p =0.001]. There was no significant main effect of victim age [F(1, 126) =0.71, p =0.40]. A significant main effect of case type [F(3, 378) =25.00, p <0.001], followed by mean comparisons with Bonferroni corrections, indicated that jurors considered the defendant to be less guilty in the stranger (M =3.21, SD =1.69) than in the incest (M =4.24, SD =1.47), day-care (M =4.72, SD =1.40), and teacher abuse (M =3.87, SD =1.73) cases, and more guilty in the day-care case than each other case (p <0.05). None of the interactions were significant (all F-values <3.64, p-values >0.05).

Next, we tested our main hypothesis, that attitudes toward children would not only differ between men and women, but would also mediate the effect of gender (dichotomized, 0 = men, 1 = women) on guilt judgments. Bootstrap mediation analyses using

| Table 2. Inter-scale correlations in Experiments 1, 2, and 3 |
|-----------------|-----------------|-----------------|-----------------|-----------------|
|                | Children's Believability | Child Victim Empathy | Opposition to Adult/Child Sex | General empathy |
| **Experiment 1 (N =132)** |                          |                          |                          |                  |
| Children's Believability |                          |                          |                          |                  |
| Session 1          | 0.64*                  | 0.44*                  | 0.33*                  | 0.39*           |
| Session 2          |                        | 0.44*                  | 0.40*                  | 0.23*           |
| Child Victim Empathy |                          |                          |                          |                  |
| Session 1          | –                      | 0.69*                  | 0.36*                  | 0.41*           |
| Session 2          | –                      |                        | 0.32*                  | 0.32*           |
| Opposition to Adult/Child Sex |                          |                          |                          |                  |
| Session 1          | –                      | –                      | –                      | 0.66*           |
| Session 2          | –                      | –                      | –                      | 0.15            |
| **Experiment 2 (N =392)** |                          |                          |                          |                  |
| Children's Believability | 0.41*                  |                          |                          | 0.24*           |
| Child Victim Empathy | –                      | 0.42*                  |                          | 0.34*           |
| Opposition to Adult/Child Sex |                          |                          |                          | 0.16*           |
| **Experiment 3 (N =475)** |                          |                          |                          |                  |
| Children's Believability | 0.41*                  |                          |                          | 0.34*           |
| Child Victim Empathy | –                      | 0.39*                  |                          |                  |

*p <0.05.
the “Indirect” SPSS macro (Preacher & Hayes, 2008) with all mediators added simultaneously (multicollinearity was not a significant concern, all $r$-values $< 0.44$) revealed that women were significantly more likely than men to feel empathy for child victims, believe children, and oppose adult/child sex (see Table 3 for all mediation results); however, only the Children’s Believability Scale mediated the effect of gender on degree of guilt: women were more likely to perceive children as believable, and in turn endorse a higher degree of guilt. The other attitudes did not significantly predict degree of guilt.

**EXPERIMENT 2**

Experiment 1 revealed that our newly constructed scales were reliable and had acceptable test–retest reliability. As predicted, women endorsed more child victim empathy, children’s believability, and opposition to adult/child sex, and believability partly explained gender differences in case judgments. Experiment 2 was a replication and extension of Experiment 1. We investigated the role of an additional attitudinal construct (attitudes toward women), case evidence strength, and a different child age comparison (5 vs. 14 years). We predicted that jurors would find a 5-year-old victim to be more credible than a 14-year-old victim, whom we expected to be above a perceived threshold age for suspicions related to dishonesty and consent issues discussed earlier. Other predictions are discussed next.

**Attitudes Toward Women**

In feminist literature, the sexual abuse of women and children is conceptualized as an outgrowth of dominant, possessive positions held by men in relation to women and children throughout history (Herman, 1981; Herman & Hirschman, 1977; Rush, 1980). Compared with women, men endorse less pro-woman attitudes (Spence, Helmreich, & Stapp, 1973; Thorton, Ryckman, & Robbins, 1982). Anti-feminist, gender-stereotyped attitudes are related to rape acceptance (Anderson, Cooper, & Okamura, 1997) and to pro-defense evaluations in mock rape trials, among women and men (Ong & Ward, 1999). Studies testing the relation between pro-women attitudes and case-related judgments have found no effects (Maynard & Wiederman, 1997) or mixed effects (Quas et al., 2002), but Gabora et al. (1991, 1993) reported that pro-women attitudes were correlated positively with jurors’ ratings of victim credibility and negatively with defendant credibility. We predicted that jurors with anti-feminist views would react less negatively than others to child sexual abuse cases, men and women would differ in their feminist views, and that these attitudes would in part explain gender differences in case judgments.

**Case Strength**

Case evidence is (and should be) the most important determinant of juror verdicts (Devine et al., 2001; Kalven & Zeisel, 1966). In child sexual abuse cases, strong evidence such as DNA and medical evidence (Bradshaw & Marks, 1990; Golding et al., 2000), defendant criminal history (Bottoms & Goodman, 1994), and credible expert evaluations (Kovera, Levy, Borgida, & Penrod, 1994) increase, and weak
<table>
<thead>
<tr>
<th>Study 1</th>
<th>Degree of guilt</th>
<th>Victim responsibility</th>
<th>Defendant responsibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Victim Empathy</td>
<td>0.15 (0.14)</td>
<td>[-0.06, 0.23]</td>
<td>N/A</td>
</tr>
<tr>
<td>Children’s Believability</td>
<td>0.51* (0.10)</td>
<td>[0.15, 0.42]*</td>
<td>N/A</td>
</tr>
<tr>
<td>Opposition to Adult/Child Sex</td>
<td>-0.16 (0.11)</td>
<td>[-0.22, 0.03]</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study 2</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Victim Empathy</td>
<td>0.18* (0.08)</td>
<td>[0.004, 0.18]*</td>
<td>-0.17* (0.06)</td>
</tr>
<tr>
<td>Children’s Believability</td>
<td>0.41* (0.05)</td>
<td>[0.18, 0.33]*</td>
<td>-0.07* (0.04)</td>
</tr>
<tr>
<td>Opposition to Adult/Child Sex</td>
<td>-0.16* (0.06)</td>
<td>[-0.11, -0.02]*</td>
<td>-0.28* (0.04)</td>
</tr>
<tr>
<td>Attitudes Toward Women</td>
<td>-0.21 (0.13)</td>
<td>[-0.22, 0.02]</td>
<td>-0.37* (0.09)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study 3</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Victim Empathy</td>
<td>-0.01 (0.09)</td>
<td>[-0.09, 0.07]</td>
<td>-0.49* (0.07)</td>
</tr>
<tr>
<td>Children’s Believability</td>
<td>0.49* (0.05)</td>
<td>[0.11, 0.27]*</td>
<td>0.21* (0.04)</td>
</tr>
<tr>
<td>Opposition to Adult/Child Sex</td>
<td>-0.03 (0.06)</td>
<td>[-0.05, 0.03]</td>
<td>-0.05 (0.05)</td>
</tr>
</tbody>
</table>

*p < 0.05.

Note: The mediation models provided estimates of the direct and indirect effects, as well as the effects of gender on each mediator and the effect of each mediator on case judgments. The table includes the estimates and standard errors for mediator effects on judgments (e.g., effect of child victim empathy on degree of guilt), as well as the indirect effects of gender on case judgments via each mediator. Significant indirect effects are indicated by confidence intervals that do not include 0. LL, lower limit of the confidence interval; UL, upper limit of the confidence interval.
evidence such as suggestive child interviewing (Castelli, Goodman, & Ghetti, 2005), decreases convictions.

We designed our cases to be weak, neutral, or strong in terms of case evidence, pre-tested to ensure this would be reflected in verdicts, and predicted that evidence strength would moderate the effects of other variables in line with Kalven and Zeisel’s, (1966) “liberation hypothesis,” which suggests that jurors’ biases influence verdicts most when the evidence is ambiguous, because ambiguous evidence liberates jurors’ from evidentiary constraints (Devine et al., 2001; Kassin & Wrightsman, 1988; but see Leippe & Romanczyk, 1987). Thus, we expected that juror gender and attitudes would have stronger effects in cases with weaker evidence.

Method

Experiment 2 reflected a 2 (juror gender: men, women) × 2 (victim age: 5 vs. 14 years) × 3 (case strength: strong, weak, balanced) factorial design, with victim age varied between subjects and case strength varied within subjects. Case type (incest, day-care abuse, stranger abduction, teacher assault) was also varied to ensure generalizability, but we collapsed across it because it did not moderate the effect of other variables in Experiment 1.

Participants

Participants were 392 introductory psychology students from the same northeastern large public university. Participant age ranged from 17 to 43 years ($M =19$ years, $SD =2.41$; 50% women, 80% White, 10% African American, 5.5% Asian American, 4.5% “other”) who received course credit for participation.

Materials

The demographic questions and attitude questionnaire were exactly the same as those employed in Experiment 1.

*Attitudes Toward Women Scale, Short Version.* (AWS; Spence et al., 1973). The 25-item short version of the AWS measures attitudes toward women and women’s roles (e.g., “Women should worry less about their rights and more about becoming good wives and mothers”) and correlates highly (0.97) with the longer version. The scale was reliable in our sample ($\alpha =0.88$).

*Case Scenarios and Judgment Scales.* Scenarios were similar to those in Experiment 1, except that each had three case strengths (strong, balanced, and weak), and the day-care center became an after-school activity center in cases involving a 14-year-old victim to maintain believability. Case strength was varied by systematic manipulation of the presence or extent of evidence, such as expert testimony, medical findings, or defendant past acts. For example, evidence strength was increased by adding the following case elements: past accusations that the defendant had sexually abused a child (teacher case); stronger quality of psychologists’ expert testimony (day-care case); defendant history of abusing his wife and child (incest case); more accurate expert medical testimony (stranger case); and less suggestive child witness interview (day-care case). Pilot-testing ensured that conviction rates varied as a function of the case strength manipulation. Conviction rates in pilot testing and the main study were
approximately 20% in the weak cases, 50% in the balanced cases, and 80% in the strong cases.

The verdict and degree-of-guilt measures were identical to those in Experiment 1; for degree of guilt, \( M = 3.71, \ SD = 1.79 \). We added separate measures of victim and defendant responsibility: “If you believe that the abuse occurred, how responsible do you think the [victim/defendant] was for the abuse?” on six-point scales ranging from 1 (not at all responsible) to 6 (extremely responsible) (victim responsibility, \( M = 1.68, \ SD = 1.17 \); defendant responsibility \( M = 5.47, \ SD = 1.16 \)).

**Procedure**

The procedure was the same as for Experiment 1, except that participants completed all measures during one session. Counter-balancing controlled for potential order-of-presentation effects: half of the participants received the attitude and empathy scales first, and half received the scenarios and judgment scales first.

**Results**

Reliability analyses revealed acceptable item-total correlations and scale internal consistencies (\( \alpha \)-values = 0.68–0.87; mean inter-item correlations = 0.17–0.35). Inter-scale correlations were significant, yet not as high as individual scale reliabilities, suggesting that the scales tapped distinct constructs (see Table 2). Women were significantly more empathic toward child victims, more believing in them, more opposed to adult/child sex, and more positive in their attitudes toward women (see Table 1).

**Case Judgments**

Verdicts were well balanced between not guilty (48.5%) and guilty (51.5%), with significantly fewer men (43%) than women (60%) voting guilty \([\chi^2 (N = 392) = 41.78, \ p < 0.001]\). Separate 2 (juror gender: men, women) × 2 (victim age: 5, 14 years) × 3 (case strength: weak, medium, strong) ANOVAs were performed on ratings of degree-of-guilt, defendant responsibility, and victim responsibility. As expected, women endorsed higher degree of guilt than did men \((M = 3.38, \ SD = 1.78 \ vs. \ M = 4.00, \ SD = 1.74) [F(1, 1533) = 63.18, \ p < 0.001]\), rated the victim as less responsible \((M = 1.58, \ SD = 1.10 \ vs. \ M = 1.77, \ SD = 1.22) [F(1, 1152) = 4.44, \ p = 0.04]\), and the defendant as more responsible \((M = 5.53, \ SD = 1.08 \ vs. \ M = 5.40, \ SD = 1.24) [F(1, 1123) = 4.10, \ p = 0.04]\). When victims were 5 years old (vs. 14 years old), they were rated as less responsible \((M = 1.43, \ SD = 1.00 \ vs. \ M = 1.92, \ SD = 1.26) [F(1, 1152) = 49.82, \ p < 0.001]\), and defendants were rated as more responsible \((M = 5.56, \ SD = 1.12 \ vs. \ M = 5.37, \ SD = 1.19) [F(1, 1124) = 10.86, \ p = 0.001]\).

There were significant main effects of case strength for degree of guilt (weak, \( M = 2.57, \ SD = 1.48 \); medium, \( M = 3.65, \ SD = 1.69 \); strong, \( M = 4.84, \ SD = 1.40) [F(1, 1533) = 295.52, \ p < 0.001]\), and defendant responsibility (weak, \( M = 5.15, \ SD = 1.48 \); medium, \( M = 5.47, \ SD = 1.15 \); strong, \( M = 5.69, \ SD = 0.82) [F(1, 1123) = 18.97, \ p < 0.001]. Planned comparisons revealed that all comparisons were significant for degree of guilt (all \( F\)-values > 80.34, \( p\)-values < 0.001) and defendant responsibility (all \( F\)-values > 3.18, \( p\)-values < 0.05). Case strength did not significantly interact with any variables.
Mediation Analyses

Bootstrap mediation analyses with the Indirect macro revealed that women were significantly more likely than men to feel empathy for child victims, believe children, oppose adult/child sex, and hold positive attitudes toward women (see Table 3). Although our mediators were significantly correlated (see Table 2), multicollinearity was not a significant concern (the highest correlation was between children’s believability and opposition to adult/child sex, \( r = 0.44 \)). Child victim empathy and children’s believability mediated the effect of gender on degree of guilt: Women felt more empathy for children and perceived them as believable, and in turn endorsed higher degree of guilt.\(^2\)

Child victim empathy, children’s believability, opposition to adult/child sex and attitudes toward women also mediated the effect of juror gender on attributions of responsibility for the victim (see Table 3): Participants with less empathy, less belief in children, less opposition to adult/child sex, and less positive attitudes toward women predictably endorsed more victim responsibility, explaining the gender differences in responsibility ratings. Finally, child victim empathy and opposition to adult/child sex also mediated the effect of juror gender on attributions of defendant responsibility: the higher the empathy and the opposition to adult/child sex, the stronger the attribution of responsibility to the defendant.

Summary

Experiment 2 further supported our theory that gender differences in child sexual abuse case judgments – not only guilt but also victim and defendant responsibility – are driven in part by individual differences in attitudes and in empathy. As in Experiment 1, the construct generally most predictive of judgments was the Children’s Believability Scale, but the Child Victim Empathy Scale also explained gender differences in guilt judgments, and all four scales (Child Victim Empathy, Children’s Believability, Opposition to Adult/Child Sex, and Attitudes Toward Women) explained gender differences in victim responsibility judgments. Victim age affected defendant and victim responsibility judgments as predicted, but not guilt judgments. More pro-prosecution judgments were rendered as evidence strength increased, yet this did not influence the extent to which extralegal factors such as gender, victim age, and attitudes influenced verdicts.

EXPERIMENT 3

In Experiment 3, we replicated and extended our first two experiments by investigating whether our findings generalized to a representative community sample of jury-eligible citizens participating at a courtroom. In many mock jury studies, participants differ

\(^2\) Surprisingly, overall, the higher the opposition to adult/child sex, the lower the degree of guilt. The zero-order correlation between Opposition to Adult/Child Sex and degree of guilt was positive as expected \((r = 0.06, p < 0.05)\), but the effect became negative once Children’s Believability Scale was added to the regression equation. This suppressor effect (see MacKinnon, Krull, & Lockwood, 2000; Paulhus, Robins, Trzesniewski, & Tracy, 2004) indicates that, after accounting for the overlap between Opposition to Adult/Child Sex and Children’s Believability, the remaining unique effect of the former is negative, although the total effect is positive (as evinced by the correlation).
from the average profile of community members in several aspects. The question debated in the field of psychology and law is whether those aspects influence legal decision-making, or even decision-making in only certain types of cases (Bornstein, 1999; Diamond, 1997; Wiener, Krauss, & Lieberman, 2011). Compared with community members called for jury duty, undergraduate students are usually younger, of relatively high SES, have higher need for cognition (i.e., the extent to which people enjoy effortful thinking; Cacioppo & Petty, 1982), and are less likely to be parents (McCabe, Krauss & Lieberman, 2010; Sears, 1986; Quas, Thompson, & Clarke-Stewart, 2005). But there is little evidence that these factors are linked to decision-making in child sexual abuse cases (e.g., McCauley & Parker, 2001). Brigham (1996) found that increasing levels of experience with children among mock jurors was associated with decreasing attribution of responsibility to a child victim and more guilt to a perpetrator, but McCauley and Parker (2001) and Goodman et al. (1998) did not uncover evidence for this.

In fact, a meta-analysis conducted by Bornstein (1999) found few differences between the judgments of undergraduates and community mock jurors, although only one of the studies addressed child sexual abuse. We know of three studies comparing undergraduates’ and community members’ perceptions of child sexual abuse cases, and those studies suggest that, although there are demographic and psychological differences between community and student samples (Quas et al., 2005), their case judgments and beliefs in child abuse cases are very similar (Crowley, O’Callaghan, & Ball, 1994; Kovera & Borgida, 1997; Quas et al., 2005).

In Experiment 3, undergraduate and community member participants completed the Children’s Believability, Empathy for Child Victims, and Opposition to Adult/Child Sex scales, read two child sexual abuse case scenarios (i.e., teacher abuse and incest), and gave case judgments. We expected to replicate the pattern of findings reported in Experiments 1 and 2 in the community member sample.

Method

Participants

There were 475 community members (57% women; $M_{age} =$38 years; 75% Caucasian, 11% African American, 5% Hispanic, 4% Asian American, 1% Native American, and 4% “other”). Only 1% had less than a high school education; the highest academic achievement of the remainder was a high school diploma or GED (13%), an associate’s degree (9%), some college (35%), bachelor’s degree (21%), some graduate school (8%), or graduate degree (13%). Forty-seven percent were employed or self-employed, 7% unemployed, 13% retired, 16% students, 10% working students, 5% homemakers, and 2% were otherwise employed. Forty-two percent were married, and 56% had at least one child. Undergraduate participant data were from Experiment 2.

Materials

All materials for the undergraduate sample were reported previously in Experiment 2. The following were used for the community sample.

Demographic Questions. A demographic questionnaire assessed age, gender, marital status, SES, ethnicity, educational background, and parenthood.
Attitude Questionnaire. This was the same questionnaire as in Experiment 2.

Case Scenarios and Judgment Scales. Community jurors were given two case scenarios: the balanced case strength and 5-year-old-victim versions of the incest and teacher assault scenarios used in Experiment 2. Measures of degree of guilt and victim and defendant responsibility were the same as in Experiment 2, degree of guilt $M=3.58$, SD =1.68; victim responsibility $M=1.56$, SD =1.18; defendant responsibility $M=5.48$, SD =1.17.

Procedure

The community participants were contacted randomly from California voter lists by a marketing agency. They arrived at a mock courtroom of a law school, received informed consent and the demographic questions, then completed a separate study in which they played the role of juror in a staged mock trial of a child witness case (for details, see Goodman et al., 2006; Redlich et al., 2002), then completed the current study by reading the two case scenarios in counterbalanced order and completing case judgment questions. Counter-balancing controlled for potential order-of-presentation effects: half of the participants received the attitude and empathy scales first, and half received the scenarios and judgment scales first. After completing this study, they were paid $25 for participating in the two studies, thanked, and debriefed.

Results

First, for the community sample only, we present the results of analyses testing: scale reliability; gender differences in attitudes, empathy, and case judgments; and the mediation of gender differences in case judgments by attitudes and empathy. Then, we present analyses directly comparing the two samples on the main dependent measures.

Attitude and Empathy Scales

Reliability analyses on the community members’ responses to multi-item Child Victim Empathy, Opposition to Adult/Child Sex, and Children’s Believability scales again revealed acceptable item-total correlations internal consistencies ($\alpha$-values ranged from 0.68 to 0.83) and mean inter-item correlations (0.17–0.29) (see APPENDIX). Individual scale means and standard deviations are presented in Table 1. As in Experiments 1 and 2, inter-scale correlations were significant, yet not as high as the individual scale reliabilities, suggesting that the scales tapped distinct constructs (see Table 2). Also women were significantly more empathic toward child victims, more believing in them, and more opposed to adult/child sex than were men.

Case Judgments

Overall, verdicts were well balanced between not guilty (55%) and guilty (45%), with significantly fewer men (39%) than women (50%) voting guilty [$\chi^2(N=475) =9.98$, $p =0.002$]. Mediation analyses using the bootstrap mediation macro Indirect revealed that the total effects of juror gender on degree of guilt ($B=-0.38$, $SE =0.11$, $p <0.001$), victim responsibility ($B=-0.25$, $SE =0.11$, $p =0.003$), and defendant responsibility ($B =0.20$, $SE =0.09$, $p =0.02$) were all significant. Given that, in all
analyses, we dichotomized gender (0 = men, 1 = women), the coefficients indicate that, compared with men, women assigned a greater degree of guilt and defendant responsibility, but less victim responsibility.

In the community sample, a propensity to believe child victims was associated with more pro-prosecution judgments and was a reliable mediator of gender effects on judgments (see Table 3). Child victim empathy significantly mediated the effect of juror gender on attribution of defendant and victim responsibility, with increased empathy resulting in less victim responsibility, but more defendant responsibility. Surprisingly, although women (compared with men) found children, overall, to be more believable and found the victim less responsible, jurors who found the victim more believable also found her more responsible. It is possible that this illustrates a suppressor effect (see MacKinnon, Krull, & Lockwood, 2000): Women regard the victim as more responsible than do men, but also tend to believe children more because they perceive children as more competent cognitively; in turn, regarding the victim as more competent could result in attributing her more responsibility for the abuse. Finally, opposition to adult/child sex had no effect on judgments.

Relation between Gender, Experience with Children, Juror Age, and Case Judgments

As expected, in the community sample, women (vs. men) had significantly more experience with children ($M = 5.00$, $SD = 1.52$ vs. $M = 4.38$, $SD = 1.88$) ($t(471) = 3.90$, $p < 0.001$), and experience with children was related to more child victim empathy ($r = 0.21$, $p < 0.001$) and more opposition to adult/child sex ($r = 0.08$, $p = 0.02$), but experience with children was unrelated to case judgments ($r$-values $< 0.06$, $p$-values $> 0.09$). Juror age was related to less opposition to adult/child sex ($r = -0.13$, $p < 0.001$), increased attributions of victim responsibility ($r = 0.07$, $p = 0.04$), and decreased attributions of defendant responsibility ($r = -0.14$, $p < 0.001$). When we controlled for age and experience with children in our mediation models, results were unchanged.

Summary

Experiment 3 revealed that our results from Experiments 1 and 2 regarding psychological determinants of gender differences generalize to a community member sample, and that the construct most predictive of judgments was the Children’s Believability Scale. To a lesser extent, child victim empathy and opposition to adult/child sex were related to judgments of victim and defendant responsibility.

DISCUSSION

The data from three separate experiments supported our main hypotheses. Differences between the judgments of men and women in child sexual abuse cases were mediated by differences in pre-trial attitudes and empathy. Results also provided answers to other important questions regarding factors that influence mock jurors’ decisions in child sexual assault cases, as well as the generalizability of these findings from undergraduate samples to community samples.
Explaining Juror Gender Differences

The gender differences observed in verdict and degree-of-guilt judgments (Experiments 1, 2, and 3) and responsibility attributions (Experiments 2 and 3) across four hypothetical cases of child sexual assault paralleled those noted in prior research (for a review, see Bottoms et al., 2007). Specifically, women made significantly more pro-victim judgments than did men, although both genders gave ratings that were overall pro-victim and anti-defendant, and there were certainly within- as well as between-group differences. Further, these gender effects were explained by underlying gender differences in child victim empathy, children’s believability, and opposition to adult/child sex (and attitudes toward women in Experiment 2). That is, the multi-item scales we created to measure the empathy and attitude constructs were internally reliable, had predictive validity, and, as hypothesized, mediated the effects of gender on case judgments. Women are more likely than men to feel more empathy for child victims, to believe these victims are telling the truth, and to oppose sexual relations between adults and children. In turn, these latent variables (and the tendency to believe children, in particular) shape guilt and responsibility judgments in child sexual abuse cases. These latent variables are also informative in understanding within-gender differences: men who were more empathic toward child victims and who held attitudes similar to those more often held by women tended to be more pro-victim; women who were less empathic and who held relevant attitudes more like men were generally less pro-victim.

The proposed constructs were predictive of case judgments to varying degrees. Children’s Believability was a consistent predictor of (and explained gender differences in) degree-of-guilt ratings, which is not surprising considering that jurors in child maltreatment cases often consider child testimony the most important evidence (Myers et al., 1999). Believability was not linked to perceptions of defendant responsibility and was a less consistent predictor of victim responsibility attributions – decreasing perceptions of victim responsibility in Experiment 2 (where the victim was either 5 or 14 years old), but increasing them in Experiment 3 (where the victim was always 5). Perceptions of cognitive competency might result in a tendency to believe that young children are capable of accurate testimony and that they are capable of understanding the situation and are therefore more likely to be responsible.

Child victim empathy was also a strong mediator of gender effects, especially attributions of responsibility. In Experiments 2 and 3, the Child Victim Empathy Scale predicted attributions of less victim responsibility and more defendant responsibility, mirroring findings from research on perceptions of adult sexual abuse (Barnett et al., 1992; Deitz et al., 1982). Thus, women are more likely than men to understand and feel emotionally close to child victims; in turn, they blame child victims less and defendants more. These gender differences in empathy might result from the internalization of gender roles, which socialize women to be caring, concerned, and child-oriented (Barnett & Sinisi, 1990; Eagly & Wood, 1991). They might also be related to women’s disproportionately higher rate of sexual victimization (e.g., Finkelhor, 1984), which might lead to a heightened tendency to identify with other victims and to be personally concerned with sexual abuse (Haegerich & Bottoms, 2000).

Opposition to adult/child sex played a relatively large role in the determination of responsibility judgments for both victim and defendant (in Study 2), but a smaller role in guilt judgments. In part, the measure of adult/child sex attitudes tapped whether a
participant considered a child a “sexual being.” It follows that considering children capable of having sexual needs would be associated with a greater tendency to assign responsibility to them for sexual contact with adults.

Although positive attitudes toward women and feminism were not strong predictors of case judgments, they were related to less responsibility attributed to a child victim—perhaps because such attitudes reflect a desire to avoid victim-blaming in general, especially those of male perpetrators. This finding replicates research investigating the relation of feminist-oriented attitudes and child abuse victim responsibility (Gabora et al., 1991), and adult rape victim responsibility (Spanos, DuBreuil, & Gwynn, 1991).

**Victim Age**

In Experiment 1, the age difference (5 or 10 years old) was not large enough to elicit a difference in case judgments, particularly because both victims were preadolescent. In Experiment 2 (5 or 14 years old), we found that, although guilt judgments were not affected by victim age, attributions of responsibility were: jurors who read about a 14-year-old (vs. a 5-year-old) victim attributed more responsibility to the victim and less to the defendant. Thus, it appears that, although adolescent victims are blamed more than are younger victims for the abuse, these attributions do not necessarily result in differences in guilt ratings—perhaps because, in cases involving children (as opposed to adult women), responsibility is unrelated to legal blame: any sexual act with a juvenile is automatically considered a crime, regardless of jurors’ attributions of blame.

**Case Strength**

Previous research has found that juror biases are most likely to enter into case decisions when case evidence is ambiguous, because strong case evidence trumps individual biases in legal decisions (i.e., the “liberation hypothesis”; Devine et al., 2001; Kalven & Zeisel, 1966; Zeisel & Diamond, 1978). We also found that case evidence was the strongest predictor of judgments. Gender effects, however, persisted even in strong cases, suggesting that certain gender-related personal biases are powerful enough to manifest themselves even when the evidence is unambiguous.

**Generalizability of Findings to Community Members**

Experiment 3 extended the generalizability of this research from an undergraduate sample to a more representative community member sample. Of importance, the main hypotheses of this study replicated with a community member sample: child sexual abuse case judgments are influenced by jurors’ gender, and individual differences in attitudes and in empathy mediate this effect.

Because most mock jury studies are conducted using an undergraduate sample, our study is valuable in pointing out the similarities (i.e., consistent gender effects on judgments and attitudes) between college and community samples. Thus, we help clarify mixed findings in the field of child sexual abuse research by adding support that research using undergraduate samples does, for the most part, generalize to research using a community member sample, a finding similar to that of Block et al. (2012), who found that student and community samples did not differ in their ability to detect
false and accurate reports made by children. Our study is also consistent with studies conducted by Crowley et al. (1994) and Isquith (1988), which found few differences between undergraduates and community members when those community members did not hold child-related careers. The issue of generalizability (i.e., from student samples to adult jury-eligible populations) in psychology and law research more broadly is also discussed in Bornstein's (1999) review of the literature, and the conclusions are similar to ours: there are few consistent and meaningful differences between student and community samples when it comes to legal judgments.

The stakes are high for science in general. At least one leading journal has begun to reject otherwise strong manuscripts solely on the basis of sample. This policy might be driven in part by intuitive bias, in part by the burgeoning availability of internet samples through Mechanical Turk and other platforms that make the collection of older samples easier, but as far as we know, it is not driven by clear evidence that there are differences between humans who are in college and those in a jury pool (or in internet samples) that significantly affect legal decision-making. It is an important empirical question for many reasons, including because it is beyond many researchers' ability to conduct experiments with community samples, or at least many such experiments, due to practical cost and time constraints. Of course, if undergraduate samples threaten the scientific integrity of the data, the studies should not be published. But if not – if there is no hard evidence that studies using undergraduate samples are misleading – then keeping such work from being published deprives the field of a great deal of research that could inform science and law.

Future Directions

Future work should investigate other variables that might also account for some of the gender differences in judgments. For example, compared with men, women score higher on measures of general “liking of children” (Barnett & Sinisi, 1990); tend to make more stable and internal attributions of causality and blame to child sexual abuse offenders (Beling, Hudson, & Ward, 2001); perceive child sexual abuse as more serious (e.g., Kovera et al., 1993; for a review, see Bottoms et al., 2007); are less likely to endorse “child sexual abuse myths” (Morison & Greene, 1992); and are more likely to be sexually abused themselves (e.g., Finkelhor, 1984; McIntyre & Widom, 2010). The latter, in particular, might lead to a heightened tendency to identify with other victims and feel more similar to them (Haegerich & Bottoms, 2000). By contrast, men are more likely than women to be charged with child sexual abuse (e.g., Goodman et al., 1992). Thus, the possibility of a false accusation of child sexual abuse may seem a more threatening reality for them than for women, which might make them feel more similar to the defendant, perhaps explaining part of the gender differences.

Methodological Considerations

There are many strengths in our methodology. For example, testing hypotheses across three experiments allowed us to gather converging evidence for our theories. Replicating findings in a community sample allows us to have confidence in the generalizability of our findings. Our written scenario methodology, using evidence from actual cases,
allowed for drawing clear cause-and-effect conclusions and for testing the generalizability of our work across four different abuse scenarios.

But in the laboratory, experimental control comes at the expense of some realism. Mock jury studies cannot fully duplicate the experience of serving on a real jury. In our study, like many others, mock jurors were not exposed to live witnesses, they knew their judgments would not have a direct impact on a defendant’s life, and they rendered individual verdicts without deliberating as a group. Even so, research reveals few differences between mock jurors’ decisions in studies using written scenarios versus more elaborate videotaped testimony (e.g., Goodman, Golding, & Haith, 1984; Goodman, Golding, Helgeson, Haith, & Michelli, 1987). Mock trial studies are therefore a reasonable first step in a line of research that should lead to more and more realistic studies that include more elaborate stimuli and deliberation (Diamond, 1997).

CONCLUSION

Our work is a significant step toward expanding the theoretical research base that not only identifies juror individual difference variables that affect case decisions, but also explains those effects in terms of underlying psychological mechanisms. In addition to the obvious contributions to the psychology and law literature, it also contributes to work on gender difference. The work also has implications for law and policy. Understanding general reactions to child sexual abuse and its victims may facilitate therapeutic preparation of child victims for reactions they may encounter inside or outside a courtroom setting. Identifying specific juror biases that influence decisions might make it possible to design jury instructions that educate jurors about the role of biases in decision making, or to develop voir dire questions that identify biased jurors before they are assigned to a jury. In fact, some have advocated screening jurors in adult rape cases based on pretrial biases such as an inability to empathize with victims or defendants (see Deitz et al., 1982). Both juror education and exclusion could help protect (a) children from revictimization at the hands of unconvicted perpetrators, and (b) innocent defendants from unjust conviction. Given the high incidence of child sexual assault cases, the special problems that arise when jurors hear such cases, and our societal obligation to assure fair trials, we cannot afford to neglect inquiry into determinants of children’s perceived credibility.

ACKNOWLEDGEMENTS

The work was supported by grants to the first author from the Sigma Xi Research Society and the American Psychology-Law Society, as well as by a grant from the Department of Health and Human Services (National Center on Child Abuse and Neglect). Research reported in this paper won an award from the American Psychology-Law Society. Many thanks are extended to Philip R. Shaver, Murray Levine, and Cathy Carter for insightful guidance and support; to Leslie Dreblat, Rachel Pearlman, and Suzanne Davis for excellent research assistance; and to Allison Redlich, Jodi A. Quas, and Jianjian Qin for assistance with Experiment 3 data collection.
REFERENCES


Krauss, D. A., McCabe, J. G., & Lieberman, J. D. (2012). Dangerously misunderstood: Representative ju-

DOI: 10.1002/bsl


APPENDIX

Coefficient Alphas, Mean Inter-item Correlations (ICs), and Corrected Item-total Correlations for Multi-item Scales

I. Child Victim Empathy Scale:

Experiment 1, S1: Alpha =0.72; IC M =0.19
Experiment 1, S2: Alpha =0.72; IC M =0.20
Experiment 2: Alpha =0.68; IC M =0.17
Experiment 3: Alpha =0.68; IC M =0.17
I can understand how someone could be emotionally scarred by a childhood incident of sexual abuse.

It makes me sad to hear about children who have been sexually molested.

If I were a member of the jury in a child sexual abuse trial, I would probably be more likely to believe the adult’s testimony than the child’s, since child sexual abuse is a charge that is difficult to defend against, even if the adult is innocent.

**I would find it easier to imagine how an adult might feel during an act of child sexual abuse than how the child victim might feel.

Testifying in court about child sexual abuse probably isn’t really painful for a child.

**I would find it easier to empathize with the shame and humiliation a child victim might feel during a child sexual abuse trial than with the feelings a child molester might have during the trial.

It’s hard for me to understand why people get so upset when they hear about children being molested.

I do not really feel sorry for a child who has to testify in court.

**I can really empathize with the helplessness a child victim might feel during a sexual assault.

*It's hard for me to imagine how an adult might feel during an act of child sexual abuse than how the child victim might feel.

**I have a lot of sympathy for child sexual assault victims.

II. Children’s General Believability Scale:

<table>
<thead>
<tr>
<th>Item-total correlations</th>
<th>Scale items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp. 1</td>
<td>Exp. 2</td>
</tr>
<tr>
<td>S1</td>
<td>S2</td>
</tr>
<tr>
<td>0.48</td>
<td>0.63</td>
</tr>
<tr>
<td>0.51</td>
<td>0.55</td>
</tr>
<tr>
<td>0.41</td>
<td>0.44</td>
</tr>
<tr>
<td>0.28</td>
<td>0.49</td>
</tr>
<tr>
<td>0.40</td>
<td>0.49</td>
</tr>
<tr>
<td>0.66</td>
<td>0.69</td>
</tr>
<tr>
<td>0.42</td>
<td>0.47</td>
</tr>
</tbody>
</table>

(Continues)
Children don’t know enough about sex to be able to make up stories of sexual abuse.

Children usually tell the truth.

Even if a mother was in a custody battle with her husband and wanted to get back at him, it would be hard for the mother to convince her child to make a false report of sexual abuse against him.

I would not want my fate to rest on the testimony of a child witness.

Children do not claim to be sexually abused just to get attention.

*Children don’t have the memory abilities to give accurate testimony in child sexual abuse trials.

*Children should not be allowed to testify in child sexual abuse trials.

*Adults can’t easily "coach" children to lie about sexual abuse.

### III. Opposition to Adult/Child Sex Scale

<table>
<thead>
<tr>
<th>Item-total correlations</th>
<th>Scale items</th>
<th>Exp. 1</th>
<th>Exp. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.51</td>
<td>0.63</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>0.36</td>
<td>0.53</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>0.55</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>0.51</td>
<td>0.63</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>0.39</td>
<td>0.19</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>0.51</td>
<td>0.68</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>0.35</td>
<td>0.28</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>0.46</td>
<td>0.49</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>0.48</td>
<td>0.46</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>0.46</td>
<td>0.55</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>0.13</td>
<td>0.43</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>0.31</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>0.43</td>
<td>0.39</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>0.27</td>
<td>0.23</td>
<td>0.35</td>
<td></td>
</tr>
</tbody>
</table>

Children sometimes fantasize about having sex with a parent or other adult.

We are naive to think it is impossible for children to enjoy sexual contact with adults.

Children are not capable of wanting sexual contact with adults.

Children do not have sexual needs.

Children who desire sexual contact with children have severe mental problems.

I believe that it is possible for a child to enjoy sexual contact with an adult, whether the child admits it or not.

It's OK for adults to look at pictures of naked children to become sexually aroused.

In some cases, children invite sexual contact with adults.

I don’t remember having sexual thoughts when I was a child.

I believe that all children secretly want to have sex with adults.

Children are not sexy.

Even if a person has sexual contact with a child, I do not think his/her life should be ruined with a prison term.

It is not natural for an adult to have a sexual interest in children.

I can’t believe that adults could want to have sexual relations with children.

*Child molesters are not a serious threat to society.

*Adults who molest children should be punished harshly.

*Laws should protect the rights of adults who desire sex with children.

**I cannot understand why an adult would use force to obtain sexual relations with a child.

Notes.

*Indicates items eventually dropped from scales to increase scale reliability.

**Denotes items that were modeled after items from Deitz’s Rape Empathy Scale.

S1, Session 1; S2, Session 2.