APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **20**: 1–22 (2006)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/acp.1301

Estimating the Effects of Misleading Information on Witness Accuracy: Can Experts Tell Jurors Something They Don't Already Know?

BRADLEY D. McAULIFF[1]* and MARGARET BULL KOVERA[2]

[1]*Department of Psychology, California State University, USA*
[2]*Department of Psychology, City University of New York, USA*

SUMMARY

This study investigated potential differences between expert and lay knowledge of factors influencing witness suggestibility. Expert psychologists ($N = 58$), jurors ($N = 157$), and jury-eligible undergraduates ($N = 220$) estimated the effects of misleading information on witness accuracy for three age groups in various conditions. Respondents possessed similar knowledge of age-related trends in suggestibility, the positive effects of a pre-misinformation warning, and the negative influence of longer delays between the event/misinformation and event/final memory test. Compared to experts, laypeople underestimated the size of suggestibility differences between age groups and lacked knowledge about how event detail centrality, witness participation, and source prestige can increase witness suggestibility. Laypeople rated themselves as being largely unfamiliar with witness suggestibility research and thought that expert testimony would be beneficial. These data shed light on the potential helpfulness of expert testimony in cases involving witness suggestibility. Copyright © 2006 John Wiley & Sons, Ltd.

Psychologists testify as experts on a variety of matters in court. Traditionally, expert psychological testimony has focused on issues involving clinical assessment (Yuille, 1989); however, social framework evidence has become increasingly common. Social or behavioural scientists constituted nearly one-quarter of all scientists testifying in U.S. criminal appellate cases involving expert testimony from 1988 to 1998 (Groscup, Penrod, Studebaker, Huss, & O'Neil, 2002). Despite some judges' willingness to allow expert testimony on matters such as eyewitness identification and gender stereotyping, the admissibility of social framework evidence remains controversial.

ADMISSIBILITY OF EXPERT EVIDENCE

Judges' determinations regarding the admissibility of expert evidence are based on legal precedent (*Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 1993; *Frye v. United States*, 1923) and evidentiary standards involving the relevance, probativeness, and helpfulness of the expert's testimony. At the federal level, these standards are codified[Q1] in the Federal

*Correspondence to: B. D. McAuliff, Department of Psychology, California State University, Northridge, 18111 Nordhoff Street, Northridge, CA 91330-8255, USA. E-mail: bradley.mcauliff@csun.edu

Rules of Evidence (2001; FRE). Evidence must be relevant to be admissible (Rule 402) and may be excluded if its probative value is substantially outweighed by its potential to unfairly prejudice, confuse or mislead the jury (Rule 403). An expert may be qualified by 'scientific, technical, or other specialized knowledge' and his or her testimony must assist jurors to understand the evidence or to determine a fact at issue in the case (Rule 702). Moreover, mirroring the language of *Daubert*, expert testimony must be based on 'facts or data' that are reliable in both their methodology and application to the facts of the case. Many state legislatures have adopted evidentiary rules containing language based on the FRE.

The present study examined the potential for social framework evidence to meet the 'helpfulness' requirement of FRE Rule 702. Specifically, we were interested in social framework evidence on 'witness suggestibility' or the effects of misleading information on the accuracy of witness memory.

WITNESS SUGGESTIBILITY

Suggestibility is the extent to which certain cognitive, social and developmental factors influence an individual's ability to encode, store, retrieve and report an event (Ceci & Bruck, 1993). A great deal of suggestibility research has focused on the effects of misleading post-event information on the accuracy of witness's reports (cf. Christiaansen, Sweeney, & Ochalek, 1983; Loftus & Greene, 1980 for misinformation studies involving eyewitness identification). Typically, participants witness or participate in an event and later answer questions about that event. Once witnesses have viewed the event, some or all of them receive misleading information that is embedded within a written summary of the event, included in the interviewer's questions, or conveyed by co-witnesses.

To measure suggestibility, researchers compare the accuracy of misled and nonmisled witnesses or make comparisons within participants. In the latter design, researchers compare the accuracy of a witness's responses to misleading versus nonmisleading questions. A witness is considered to be suggestible if the witness erroneously incorporates the suggested misinformation into his or her report of the original event. Evidence for such an effect would consist of decreased accuracy on the final memory test and increased reliance on the misleading information. Various theoretical explanations for witnesses' decreased post-misinformation accuracy exist in the scientific literature (see, e.g. Belli, Windschitl, McCarthy, & Winfrey, 1992; Lindsay & Johnson, 1989; Loftus, Miller, & Burns, 1978; McCloskey & Zaragoza, 1985).

Previous witness suggestibility research has shown that factors such as witness age (Cohen & Harnick, 1980; Coxon & Valentine, 1997), event detail centrality (Heath & Erickson, 1998; Warren & Lane, 1995; Wright & Stroud, 1998), level of witness participation (Rudy & Goodman, 1991; Tobey & Goodman, 1992), warning (Christiaansen & Ochalek, 1983; Greene, Flynn, & Loftus, 1982; Saywitz & Moan-Hardie, 1994; Warren, Hulse-Trotter, & Tubbs, 1991), source prestige (Ceci, Ross, & Toglia, 1987; Dodd & Bradshaw, 1980; Lampinen & Smith, 1995; Smith & Ellsworth, 1987) and retention interval length (Belli et al., 1992; Loftus et al., 1978; Saywitz, Goodman, Nicholas, & Moan, 1991; Tucker, Mertin, & Luszcz, 1990) can moderate the effects of misleading information on witness accuracy.

DOI: 10.1002/acp

EXPERT TESTIMONY ON WITNESS SUGGESTIBILITY

Social scientific interest in witness suggestibility has spilled over into the legal system as experts have sought to impart their knowledge to jurors. Judges presiding over cases in which witness suggestibility is an issue confront the arduous task of deciding whether to admit expert testimony. Judicial uncertainty regarding whether expert testimony can assist triers of fact to understand the evidence or to determine a fact at issue has created dissent within the legal system. Some judges have admitted witness suggestibility expert testimony on the grounds that it is helpful or exceeds jurors' common understanding (*Barlow v. State*, 1998; *State of Ohio v. Gersin*, 1996). Others have refused to admit it based on the belief that jurors are sufficiently knowledgeable about the topic (*People v. Timothy Johnston*, 2000; *People v. Behrooz Kanani*, 2000; *Wright v. State*, 1998).

Can experts tell jurors something about witness suggestibility they don't already know? To answer this question, we asked expert psychologists, jurors and college students to estimate the effects of misleading information on witness accuracy for three age groups in various witnessing conditions and compared their responses. Respondents also rated their familiarity with the witness suggestibility literature and indicated their views of the appropriateness and helpfulness of expert testimony on witness suggestibility.

MEASUREMENT OF EXPERT AND JUROR KNOWLEDGE

Other researchers have measured expert and juror knowledge about psychological phenomena (see, e.g. Frazier & Borgida, 1988; Kassin, Ellsworth, & Smith, 1989; Kovera & Borgida, 1997; Morison & Greene, 1992). Those surveys contained statements describing findings in a particular area of social scientific inquiry. Respondents rated their level of agreement with each statement using a categorical agree/disagree/don't know scale or a continuous Likert-type scale. We opted for a slightly modified procedure based in part on the work of Swim (1994). Swim designed a scale that enabled participants to estimate the mean difference (i.e. effect size) between men and women for various characteristics such as friendliness and willingness to assist others. Swim evaluated the accuracy of gender stereotypes by comparing participants' effect size estimates to meta-analytic effect size estimates derived from the literature. To our knowledge, there have been no formal reliability or validity assessments of Swim's effect size estimate scale. However, Swim used two different methods of collecting effect size estimates from participants that correlated highly ($rs = 0.79$ and 0.78) with actual meta-analytic effect sizes from the literature.

Modifying Swim's scale enabled us to maximize the information we obtained from respondents. Experts and laypeople were not limited to stating whether they 'agreed, disagreed or did not know' about a particular finding, nor were they limited to the forced-choice, Likert-scale format. Instead, they were able to indicate both the *magnitude* and *direction* of the mean differences between misled and nonmisled witnesses in various conditions. They also provided effect size estimates for three different age groups as opposed to one collective estimate.

OVERVIEW

Several specific hypotheses flow from our general research question of whether experts and laypeople differ in their knowledge of factors that influence witness suggestibility. First, we

DOI: 10.1002/acp

wanted to determine whether lay knowledge of age-related differences in suggestibility was consistent with that of experts. We hypothesized that experts would indicate that preschoolers are more suggestible than older children and adults, but would not differentiate between the two older age groups, mirroring Ceci and Bruck's (1993) qualitative review of the suggestibility literature.[1] In contrast, we hypothesized that laypeople would believe that suggestibility decreases with age and that their estimates would differ for each age group as a result. We did not expect juror and college student estimates to differ based on a review of the jury simulation paradigm (Bornstein, 1999), which noted few differences in studies comparing college students and community members. Evidence supporting these predictions would consist of a significant Respondent Group × Witness Age interaction in which expert and lay estimates would vary as a function of witness age across the witnessing conditions.

Second, we sought to determine respondents' ledge of certain factors that moderate the effects of misleading information on witness accuracy and hypothesized that experts' estimates would differ from those of laypeople. We predicted experts would recognize that witnesses who were misled about central aspects of the event, actively participated in the event, were warned about the possibility of misinformation, were misled by a low prestige source, or experienced shorter delays between the event/misinformation and event/final memory test would be more accurate than witnesses who were misled about peripheral aspects of the event, passively observed the event, were not warned about the possibility of misinformation, received the misinformation from a highly prestigious source, or who experienced longer delays between the event/misinformation and event/final memory test. We believed jurors and college students would lack such knowledge. A significant Respondent Group × Witnessing Condition interaction would support this hypothesis.

Third, we wanted to determine respondents' familiarity with the witness suggestibility literature and their views about the appropriateness and helpfulness of expert testimony on witness suggestibility. We hypothesized that jurors and college students would report being less familiar with witness suggestibility research than experts, but that all respondents would view expert testimony as appropriate and beneficial to jurors. Support for these hypotheses would consist of a Respondent Group main effect for the familiarity ratings and a nonsignificant Respondent Group main effect on the appropriateness and helpfulness ratings.

METHOD

Survey instrument

We constructed an eight-page survey that consisted of the following six witnessing condition pairs or sets: (1) the to-be-remembered information was a *central* versus *peripheral* event detail; (2) the witness *participated* versus *observed* the event; (3) the witness was *warned* versus was *not warned* about the possibility of misinformation; (4) the source of the misinformation was *low* versus *high in prestige* and *authority* (5) the *delay*

[1]Even though 12 years have passed since Ceci and Bruck (1993) was published in *Psychological Bulletin*, we based our hypothesis on this piece because it remains the most authoritative, comprehensive review of age trends in suggestibility to date. One obvious limitation of this decision is that Ceci and Bruck's qualitative review does not include more recent research and developments in the scientific literature, particularly the increased emphasis on individual differences in suggestibility (see Bruck & Melnyk, 2005).

DOI: 10.1002/acp

Table 1. Witnessing condition survey items

1.	When witnesses *actively participate* in an event (e.g. playing a game with a researcher)...
2.	When witnesses *passively observe* an event (e.g. viewing a videotaped crime simulation)...
3.	When witnesses *receive a warning* that they may be misled during a study...
4.	When witnesses *do not receive a warning* that they may be misled during a study...
5.	When the source of the misleading is *high in prestige or authority* compared to the participant (e.g. the experimenter or a police officer)...
6.	When the source of the misleading or inaccurate information could be considered *equal or lower in prestige or authority* compared to the witness (e.g. same- or younger-aged peer)...
7.	When the misleading or inaccurate information pertains to a *central aspect of the event* (i.e. relating to what the bank robber did or wore)...
8.	When the misleading or inaccurate information pertains to a *peripheral aspect of the event* (i.e. relating to what other witnesses did or wore)...
9.	When the witness is misinformed or receives inaccurate information from the experimenter *within one day of witnessing the event*...
10.	When the witness is misinformed or receives inaccurate information from the experimenter approximately *1 week after witnessing the event*...
11.	When the witness is misinformed or receives inaccurate information from the experimenter approximately *1 year after witnessing the event*...
12.	When the interval between the participant witnessing the event and subsequently being questioned about the event is approximately *one day*...
13.	When the interval between the participant witnessing the event and subsequently being questioned about the event is approximately *1 week*...
14.	When interval between the participant witnessing the event and subsequently being questioned about the event is approximately *1 year*...

Note: All item stems ended with the statement 'believe that the difference in the accuracy of misled and nonmisled witnesses is'. Item numbers and order are identical to those of original survey.

between the event and the introduction of misinformation was *1 day* versus *1 week* versus *1 year*; and (6) the *delay* between the event and the final memory test was *1 day* versus *1 week* versus *1 year* (see Table 1). For each witnessing condition item, respondents estimated the difference in accuracy of misled and nonmisled witnesses for three age groups: preschoolers (5 years old and younger), elementary/middle school children (ages 6–13 years), and high school students/adults (ages 14 and up). These developmental categories were based on Ceci and Bruck's (1993) comparison of preschoolers, older children and adults.[2]

The survey included a bipolar scale to assist respondents in determining the magnitude (zero, small, moderate or large) and direction (positive, neutral and negative) of their effect size estimates (see Appendix A). The numeric labels on the scale corresponded to Cohen's (1988) guidelines for interpreting the magnitude of a particular effect size, namely the *d* statistic. According to Cohen, a *d* of 0.20 is small, a *d* of 0.50 is moderate, and a *d* of 0.80 is large. Negative values indicated that misled witnesses are more accurate than nonmisled witnesses. Positive values indicated that nonmisled witnesses are more accurate than misled witnesses. The midpoint of the scale was a neutral zero indicating no difference in the accuracy of misled and nonmisled witnesses or an effect size of $d = 0$. We provided two

[2]Because each developmental category contained witnesses of different ages and the size of the age range for each developmental category varied, we cannot know for certain that respondents were thinking of witnesses who were identical in age when responding to a specific developmental category.

DOI: 10.1002/acp

examples of how to properly use the scale and varied the magnitude and direction of misinformation effects in the examples to minimize potential response bias.

The survey also included demographic questions. We asked experts questions about their educational background, publication record and court experience. We asked jurors questions about their occupation, income, history of jury service, whether they had any children, and the current ages of those children. Finally, we asked college students questions concerning their college major, year in school, and whether they had taken classes and which professors had discussed witness suggestibility. Respondents estimated their familiarity with the witness suggestibility literature and their beliefs about the appropriateness and helpfulness of expert testimony on witness suggestibility using a series of 7-point Likert-type scales. Higher numbers indicated increased familiarity with the literature and more positive views about expert testimony. Respondents also rated on a 7-point Likert-type scale how easy or difficult it was to use the suggestibility scale. Higher numbers indicated that respondents found the scale difficult to use.

Two additional questions appeared at the end of the college survey that were not used with experts or jurors. Those questions assessed whether laypeople understood and were able to use the suggestibility scale (see Appendix B). Ninety-two per cent of the college students who completed the test questions provided correct responses. A second measure of whether laypeople had difficulty using the suggestibility scale was their responses to the difficulty item included in the survey. College students and jurors did not find the scale extremely difficult or easy to use ($Ms = 4.16$ and 4.18, respectively). Based on these indices, it appears that laypeople understood the suggestibility scale and used it appropriately to express their beliefs.

Respondents

Experts

We searched the PsychLit and Dissertation Abstracts research databases using the key words 'suggestibility' and 'misinformation' to identify psychologists who had published research in the area of witness suggestibility from 1970 to 1998. We also reviewed various conference programmes for relevant presentations and examined the references of several comprehensive reviews of the suggestibility literature. Finally, we contacted all the psychological researchers and scholars who signed the *amicus curiae* brief submitted by the Committee of Concerned Social Scientists in *State of New Jersey v. Margaret Kelly Michaels* (1994).

We sent surveys to 117 psychologists and 78 returned their surveys for an overall response rate of 67%. Fifty-eight (50%) respondents completed the entire survey and 20 (17%) returned incomplete or partially incomplete surveys. Respondents who did not complete the entire survey generally fell into two categories: those who did not consider themselves experts on witness suggestibility ($n = 13$, 11%) and those who disliked survey format ($n = 7$, 6%). After we excluded the responses of the 13 self-reported nonexperts, the final response rate to our survey was 56% (58/104).

Eight of the twenty respondents who returned incomplete surveys filled out the demographic section only. A series of one-way ANOVAs on those items revealed that complete versus incomplete respondents did not differ in age, number of publications, number of times they were asked to testify in court, number of times they had agreed to testify, or number of times they actually testified at trial. However, experts who returned incomplete surveys rated expert testimony on witness suggestibility as less beneficial to

DOI: 10.1002/acp

jurors compared to those who completed the entire survey, $F(1, 63) = 16.97$, $p \leq 0.001$, partial $\eta^2 = 0.21$, ($Ms = 3.57$ and 5.71, respectively). Similarly, experts who returned incomplete surveys also found the scale more difficult to use than those who completed the entire survey, $F(1, 59) = 4.99$, $p \leq 0.05$, partial $\eta^2 = 0.08$, ($Ms = 7.00$ and 5.21, respectively).

The majority of expert respondents ranged in age from 26 to 55 years ($M = 46$ years), were male (64%), had received a doctor of philosophy degree (97%), and specialized in cognitive/experimental psychology (48%). Eighty per cent had conducted empirical research on witness suggestibility and had included preschoolers, older children and adults in their research an approximately equal number of times across studies. Ninety-six per cent of experts stated they had published empirical research and the majority had published in scientific journals or book chapters. In total, these 51 researchers estimated they had published over 825 articles, 618 (75%) of which appeared in scientific journals, 8 (1%) of which appeared in law reviews, 192 (11%) of which appeared in books or book chapters and seven (1%) of which appeared in newsletters.

Table 2 reflects the court-related experience of experts in this study. Over half of responding experts had been asked and had agreed to testify in at least one case involving witness suggestibility; however, less than half of all experts had actually testified at trial. A few select experts reported extraordinarily high values on these three survey items so the median number of times experts were asked ($Mdn = 5$), agreed ($Mdn = 6$), and actually testified ($Mdn = 4$) may describe the court-related experience of expert respondents more accurately.

In what ways did psychologists who *had been asked* to provide expert testimony differ from those who *had not been asked* to testify? Similarly, in what ways did experts who *had agreed* to testify in court differ from those who *had not agreed* to testify? To answer these questions, we conducted one-way ANOVAs and chi-square analyses comparing each subgroup pair on a number of dimensions, including publication record, status as a member of the American Psychology-Law Society (APA's Division 41), self-reported judgements of familiarity with the suggestibility literature, and beliefs about the appropriateness and usefulness of expert testimony on witness suggestibility.

Researchers who had been asked to testify reported being more familiar with the suggestibility literature than those not asked to testify, $F(1, 53) = 26.69$, $p \leq 0.001$, partial $\eta^2 = 0.34$, ($Ms = 6.50$ and 5.41, respectively). The difference between the reported number of publications for each group was not statistically significant. Psychologists who had agreed to testify as experts believed witness suggestibility testimony was more appropriate and beneficial than did those who had not agreed to testify [Appropriateness $F(1,$

Table 2. Self-reported frequency with which experts were asked, had agreed and actually testified in court ($N = 58$)

Number of times	Experts asked to testify *N* (%)	Experts agreed to testify *N* (%)	Experts actually testified *N* (%)
1–10	17 (29)	21 (36)	18 (31)
11–35	12 (21)	9 (16)	5 (9)
36–75	3 (5)	0 (0)	0 (0)
76–300	5 (9)	4 (7)	3 (5)
301–1000	1 (2)	1 (2)	1 (2)
Total	38 (66)	35 (60)	27 (47)

DOI: 10.1002/acp

37) = 15.79, $p \leq 0.001$, partial $\eta^2 = 0.30$, ($Ms = 6.23$ and 4.25, respectively); Beneficial $F(1, 37) = 10.05$, $p \leq 0.01$, partial $\eta^2 = 0.21$, ($Ms = 5.89$ and 3.75, respectively)]. It is unclear whether these varying attitudes contributed to or were the result of psychologists' decisions to testify.

Jurors

One hundred fifty-seven US citizens reporting for jury duty in South Florida participated in our study in exchange for a $5.00 meal voucher at a nearby Burger King. Jurors averaged 53 years of age and most were white (67%), married (56%), held professional/technical or sales occupations (40%), and reported a gross family income of $60,000 or less (77%). Sixty-two per cent of jurors reported having children, 24% of whom had preschoolers (0–5 years old), 34% of whom had older children (6–13 years old), and 59% of whom had high school/adult children (14 years and older). Nearly half (47%) of participants indicated that their highest level of education was a high school diploma or college degree. Most jurors (69%) had not served on a jury before and a higher percentage of jurors who had served before decided criminal (65%) rather than civil (42%) cases. An equal number of men ($n = 77$) and women ($n = 77$) participated in the study. Three jurors elected not to complete the sex item on the demographic portion of the survey.

College students

The college student sample consisted of 220 jury-eligible undergraduate students at a public university in the southeastern United States. Undergraduates earned extra-credit for psychology classes in exchange for their participation. In accordance with the state's juror eligibility requirements at the time, all participants were US citizens, 18 years or older, and possessed a valid driver's license.

Jury-eligible students tended to be female (73%), freshmen (55%) and averaged 23 years of age. A series of one-way ANOVAs revealed that students who had taken classes covering suggestibility (60%) described themselves as being more familiar with the suggestibility literature than and those who had not (40%), $F(1, 217) = 44.52$, $p \leq 0.001$, partial $\eta^2 = 0.17$, ($Ms = 3.24$ and 1.90, respectively) and also indicated stronger beliefs about the appropriateness of expert testimony on this topic, $F(1, 215) = 5.79$, $p \leq 0.05$, partial $\eta^2 = 0.03$, ($Ms = 5.08$ and 4.68, respectively). Despite these differences, the overall effect size estimates provided by these two groups did not differ at $p \leq 0.05$. Consequently, data from both groups were included in our final college student sample.

Procedure

Experts

We used Dillman's (1978; Salant & Dillman, 1994) survey procedure to increase our response rate. The first mailing consisted of a survey and cover letter describing the study's purpose and highlighting its potential contribution to the field of legal psychology. We asked survey recipients who did not wish to participate in the study to please return their incomplete surveys with a brief explanation. Experts who elected to participate were assured that their survey responses would be anonymous. Two weeks later experts received reminder postcards encouraging them to return completed surveys or to contact us if they had not yet received a survey. Five weeks after the original mailing, experts received another survey and a final reminder letter stressing the importance of returning their

DOI: 10.1002/acp

completed surveys in order for our findings to truly reflect experts' knowledge about witness suggestibility.

Jurors

Citizens reporting for jury duty were invited to participate in a voluntary study on beliefs about witness memory. Approximately 80% of those who heard the announcement reported to a lounge adjacent to the jury assembly room for participation. After providing informed consent, jurors received the survey along with introductory instructions explaining the general purpose of witness suggestibility research and relevant terminology (e.g. 'misled' and 'nonmisled'). Jurors did not confer with one another during the study and were debriefed after completing the survey.

College students

The only procedural differences for the expert and layperson samples were first that experts filled out their surveys alone, whereas jurors and college students independently filled out their surveys in the presence of 5–10 others. Second, because the primary criterion for inclusion in the expert sample was notable expertise in the field of witness suggestibility, we did not provide the expert sample with the introductory instructions described above.

SURVEY RESULTS

Data analytic strategy

Three groups of respondents (experts, jurors and college students) estimated the effects of misleading information on witness accuracy for three age groups (preschoolers, older children, adults) repeatedly across six witnessing condition pairs or sets (central vs. peripheral event detail, participate vs. observe, warning vs. no warning, low vs. high source prestige, event/misinformation delay of 1 day vs. week vs. year, and event/final memory test delay of 1 day vs. week vs. year). We subjected respondents' effect size estimates for the 14 witnessing conditions to a multivariate analysis of variance (MANOVA) to explore differences in expert and lay knowledge of witness suggestibility as a function of witness age and witnessing condition. Respondent Group was the between-subjects independent variable and Witness Age and Witnessing Condition were the within-subjects independent variables.[3] We used the Pillai's criterion multivariate statistic to test the significance of main effects and interactions involving these three variables. We followed-up all significant multivariate main effects and interactions involving the Witnessing Condition variable with univariate F-tests and used the Bonferroni-type adjustment of $\alpha = 1 - (1 - \alpha_1)(1 - \alpha_2) \ldots (1 - \alpha_p)$ as recommended by Tabachnick and Fidel (2007) so that the overall alpha level for the set of Witnessing Condition univariate tests was equal to 0.05 ($\alpha = 0.008$ for each univariate test). Simple main effects tests and simple comparisons were conducted when appropriate. Unless otherwise indicated, the mean differences reported met traditional levels of statistical significance ($p \leq 0.05$).

[3]One reviewer recommended that we include Respondent Gender as an additional between-subjects variable in light of the gender effects observed by Quas et al. (2005) who found that women were more accurate than men when asked to predict the outcomes of several well-known studies of children's suggestibility. We did and neither the main effect nor any of the interactions involving Respondent Gender were statistically significant.

DOI: 10.1002/acp

Table 3. Respondents' suggestibility effect size estimates for different witness age groups

Witness age group	Respondent group		
	Experts *M* (*SD*)	Jurors *M* (*SD*)	Students *M* (*SD*)
Preschoolers	0.59[a] (0.34)	0.50[a] (0.29)	0.52[a] (0.42)
Older children	0.39[b] (0.20)	0.42[b] (0.24)	0.44[b] (0.30)
Adults	0.29[c] (0.20)	0.31[c] (0.26)	0.38[c] (0.27)

Note: Means with unique superscripts within each Respondent group differed at the $p \leq 0.05$ level.

Respondents' suggestibility effect size estimates

A 3 Respondent Group × 3 Witness Age × 14 Witnessing Condition MANOVA revealed a significant main effect for Witness Age, Mult. $F(2, 422) = 72.72$, $p \leq 0.001$, partial $\eta^2 = 0.26$ and Witnessing Condition, Mult. $F(13, 411) = 10.31$, $p \leq 0.001$, partial $\eta^2 = 0.25$. However, both main effects were qualified by a significant interaction with Respondent Group. The Respondent Group × Witness Age interaction, Mult. $F(4, 846) = 8.44$, $p \leq 0.001$, partial $\eta^2 = 0.04$ and Respondent Group × Witnessing Condition interaction, Mult. $F(26, 824) = 1.66$, $p \leq 0.05$, partial $\eta^2 = 0.05$ were significant. The Respondent Group main effect, Witnessing Condition × Age interaction, and the Respondent Group × Witness Age × Witnessing Condition interaction were nonsignificant, Mult. $F(2, 423) = 1.28$, $p =$ n.s., partial $\eta^2 = 0.01$, Mult. $F(26, 398) = 0.63$, $p =$ n.s., partial $\eta^2 = 0.04$, and Mult. $F(52, 798) = 0.79$, $p =$ n.s. partial $\eta^2 = 0.05$, respectively.

Age-related differences in witness suggestibility

Because the Respondent Group × Witness Age interaction did not vary as a function of Witnessing Condition, we conducted follow-up tests to determine if the MANOVA results supported our first hypothesis that the pattern of age effects in suggestibility would vary for experts compared to laypeople (but not for jurors compared to college students) across all witnessing conditions. Simple main effect tests and simple comparisons revealed the effect size estimates for each Witness Age group within each Respondent Group differed at statistically significant levels (see Table 3). All three respondent groups shared the belief that suggestibility decreases with age as evidenced by their larger effect size estimates for preschoolers and successively smaller effect size estimates for older children and adults.

We were also interested in whether the relative size of differences between certain witness age groups varied across respondents. In other words, did experts and laypeople vary in how large or small they estimated the suggestibility differences to be between preschoolers versus older children, preschoolers versus adults, and older children versus adults? To answer this question, we calculated each respondent's difference score for the three witness age group comparisons and subjected the data to one-way Respondent Group ANOVAs.

Analyses revealed that size of difference between preschoolers versus older children and preschoolers versus adults varied for experts compared to laypeople, but not for jurors compared to college students, $F(2, 432) = 10.05$, $p \leq 0.001$, partial $\eta^2 = 0.04$ and $F(2, 432) = 6.31$, $p \leq 0.01$, partial $\eta^2 = 0.03$, respectively. Compared to experts, jurors and college students underestimated the size of the suggestibility differences between preschoolers versus older children (Ms = 0.20, 0.08 and 0.08, respectively) and preschoolers versus adults (Ms = 0.31, 0.20 and 0.14, respectively). Analysis of

DOI: 10.1002/acp

respondents' difference scores for the older children versus adults comparison revealed agreement between experts and jurors, but not college students, $F(2, 432) = 4.32$, $p \leq 0.05$, partial $\eta^2 = 0.02$. Compared to experts and jurors, college students underestimated the size of suggestibility differences between older children versus adults (Ms = 0.11, 0.11, and 0.06, respectively). A series of interaction contrasts for respondents' difference scores yielded identical results.

Witnessing condition moderators of suggestibility

To follow-up the significant Respondent Group × Witnessing Condition interaction, we conducted six repeated measures ANOVAs that compared respondents' effect size estimates for the various witnessing condition pairs or sets. These analyses evaluated our second hypothesis that experts would be more sensitive to the moderators of the misinformation/accuracy relation than laypeople.

First we examined whether respondents believed that receiving misleading information about central versus peripheral crime details influenced witness suggestibility. A 3 Respondent Group × 2 Event Detail Centrality repeated measures ANOVA revealed a significant main effect for Event Detail Centrality, $F(1, 432) = 46.03$, $p \leq 0.001$, partial $\eta^2 = 0.10$. That main effect was qualified by a significant Respondent Group × Event Detail Centrality interaction, $F(2, 432) = 7.15$, $p \leq 0.001$, partial $\eta^2 = 0.03$ (see Table 4). Only experts indicated that witnesses were better able to resist misleading information concerning details central to the event (as evidenced by smaller effect size estimates) than misinformation involving peripheral event details. The difference between the central and peripheral event detail conditions for jurors and college students was not statistically significant.

Next we examined whether experts and laypeople believed that individuals have more impervious memories for events they actively participated in versus passively observed. A 3 Respondent Group × 2 Witness Participation repeated measures ANOVA revealed a significant main effect for Witness Participation, $F(1, 431) = 11.09$, $p \leq 0.001$, partial $\eta^2 = 0.03$. The main effect was qualified by a significant Respondent Group × Witness Participation interaction, $F(2, 431) = 6.73$, $p \leq 0.001$, partial $\eta^2 = 0.03$ (see Table 4). Experts reported that witnesses who actively participated in the to-be-remembered event were more resistant to suggestion than those who passively observed the event. There was no statistically significant level of participation differences for jurors and college students.

Do experts and laypeople believe that warning witnesses about the possibility of impending misinformation inoculates them against its detrimental effects? A 3 Respondent Group × 2 Warning repeated measures ANOVA revealed a significant main effect for Warning, $F(1, 432) = 24.38$, $p \leq 0.001$, partial $\eta^2 = 0.05$, such that respondents reported that nonwarned witnesses were more easily misled ($M = 0.41$) than warned witnesses ($M = 0.33$). The Respondent Group × Warning interaction was nonsignificant, $F(2, 432) = 2.31$, $p =$ n.s., partial $\eta^2 = 0.01$. Experts and laypeople agreed that nonwarned witnesses were more susceptible to the negative effects of misleading information than warned witnesses (see Table 4).

We also examined the influence of receiving misleading information from a source that is either low or high in prestige on respondents' effect size estimates. A 3 Respondent Group × 2 Source Prestige repeated measures ANOVA revealed a significant main effect for Source Prestige, $F(1, 432) = 62.38$, $p \leq 0.001$, partial $\eta^2 = 0.13$. That main effect was qualified by a significant Respondent Group × Source Prestige interaction, $F(2, 432) = 3.08$, $p \leq 0.05$, partial $\eta^2 = 0.01$ (see Table 4). Experts reported that misinformation

DOI: 10.1002/acp

Table 4. Respondents' suggestibility effect size estimates for different Witnessing Conditions

Witnessing Condition	Respondent group		
	Experts *M* (*SD*)	Jurors *M* (*SD*)	Students *M* (*SD*)
Event detail centrality			
Central	0.29[a] (0.25)	0.36 (0.30)	0.40 (0.38)
Peripheral	0.51[b] (0.29)	0.43 (0.29)	0.47 (0.37)
Level of witness participation			
Participate	0.31[a] (0.21)	0.38 (0.32)	0.43 (0.23)
Observe	0.46[b] (0.26)	0.39 (0.31)	0.43 (0.31)
Warning			
Present	0.26 (0.21)	0.30 (0.33)	0.36 (0.29)
Absent	0.42 (0.25)	0.41 (0.32)	0.41 (0.35)
Source prestige			
Low or Equal Status	0.29[a] (0.17)	0.34 (0.29)	0.34 (0.31)
High Status	0.54[b] (0.34)	0.46 (0.36)	0.46 (0.43)
Event/misinformation delay			
1 Day	0.41 (0.24)	0.42 (0.30)	0.44 (0.40)
1 Week	0.46 (0.25)	0.46 (0.34)	0.51 (0.52)
1 Year	0.55 (0.31)	0.50 (0.39)	0.59 (0.94)
Event/final memory test delay			
1 Day	0.37 (0.24)	0.34 (0.30)	0.41 (0.33)
1 Week	0.47 (0.31)	0.41 (0.29)	0.46 (0.41)
1 Year	0.56 (0.33)	0.52 (0.35)	0.57 (0.69)

Note: Means with unique superscripts within each Witnessing Condition variable differed at the $p \leq 0.05$ level.

provided by a source high in prestige and authority had a stronger, more negative impact on witness accuracy compared to when the misinformation came from an equal or lesser status source than the witness. In contrast, lay effect size estimates for the high versus low source prestige conditions did not significantly differ.

The final two moderators dealt with the potential effect of time delays between the event/misinformation and the event/final memory test. First, a 3 Respondent Group × 3 Event/Misinformation Delay repeated measures ANOVA revealed that as the amount of time that elapsed between the to-be-remembered event and the introduction of misinformation increased, all respondents reported that the misinformation exerted a more powerful negative effect on witness accuracy, $F(2, 428) = 6.83$, $p \leq 0.001$, partial $\eta^2 = 0.03$ (Day $M = 0.43$, Week $M = 0.49$, and Year $M = 0.56$). Follow-up tests revealed that only the effect sizes reported for the day and year delay differed at the $p \leq 0.05$ level. The main effect for Event/Misinformation Delay was not qualified by any other significant interactions.

The last moderator we explored was the time interval between exposure to the event and the final memory test. Similar to the results of the Event/Misinformation Delay ANOVA, a 3 Respondent Group × 3 Event/Final Memory Test Delay repeated measures ANOVA revealed that respondents' effect size estimates increased as the time interval between the event and the final memory test increased, $F(2, 424) = 20.26$, $p \leq 0.001$, partial $\eta^2 = 0.09$. Respondents provided larger effect size estimates when the memory test occurred 1 year after the event ($M = 0.55$) than when it occurred 1 week ($M = 0.44$) or day ($M = 0.38$) after the event. The main effect for Event/Final Memory Test Delay was not qualified by any other significant interactions.

DOI: 10.1002/acp

Familiarity, appropriateness and helpfulness of expert testimony

Jurors rated themselves as being less familiar with the suggestibility literature than did college students, who rated themselves as being less familiar than did experts, $F(2, 431) = 150.56$, $p \leq 0.001$, partial $\eta^2 = 0.41$, ($Ms = 2.24$, 2.70, and 6.16 respectively). Similar findings emerged for respondents' ratings of the appropriateness of expert testimony on witness suggestibility. Jurors were more neutral regarding the appropriateness of expert testimony than were college students and experts, $F(2, 427) = 37.25$, $p \leq 0.001$, $\eta^2 = 0.41$, ($Ms = 4.26$, 4.93 and 6.05, respectively). College students thought that expert testimony was more appropriate than did jurors but less appropriate than did experts. Finally, although all respondents thought that expert testimony on witness suggestibility would be beneficial to jurors in actual cases, jurors did to a lesser degree than college students and experts, whose responses did not differ significantly, $F(2, 428) = 5.20$, $p \leq 0.05$, $\eta^2 = 0.02$, ($Ms = 5.08$, 5.49 and 5.71, respectively).

DISCUSSION

We generated several specific hypotheses to evaluate whether experts and laypeople differ in their knowledge of factors that influence witness suggestibility. First, we predicted experts would indicate that preschoolers are more suggestible than older children and adults, who are equally suggestible, and that lay estimates of suggestibility would decrease as witness age increased. Second, we predicted that experts would be more knowledgeable about the moderating effects of certain variables compared to laypeople. Third, we expected jurors to rate themselves as being the least familiar with the witness suggestibility literature compared to experts and college students, but that all three respondent groups would rate expert testimony as being appropriate and helpful to jurors in actual cases. We discuss the results for each of these hypotheses in turn.

Age-related differences in suggestibility

According to Cohen's (1988) guidelines for interpreting the *d* statistic, experts and laypeople estimated that misleading information had a moderate effect on preschoolers' accuracy (*d*s ranged from 0.50 to 0.59) and a small to moderate effect on the accuracy of older children (*d*s ranged from 0.39 to 0.44) and adults (*d*s ranged from 0.29 to 0.38), and their effect size estimates differed significantly for each age level. These results failed to support our hypothesis that experts' effect size estimates would mirror Ceci and Bruck's (1993) conclusion that preschoolers are more suggestible than older children or adults, who are equally suggestible. These findings did support our hypotheses that lay estimates would differ for all three age groups (with adults perceived as least suggestible) and that juror and college student estimates would not differ.

These results are consistent with previous research investigating individuals' perceptions of age-related differences in suggestibility. Earlier studies have found that college students (Quas, Thompson, & Clarke-Stewart, 2005; Ross, Dunning, Toglia, & Ceci, 1990, Experiment 3), jurors (Quas et al., 2005), judges (Cashmore & Bussey, 1996), and other legal professionals (Brigham & Spier, 1992; Leippe, Brigham, Cousins, & Romanczyk, 1989) believe that younger children are more susceptible to misinformation than adults. Ninety-four per cent of eyewitness experts responding to a recent survey indicated that the research finding that 'younger children are more vulnerable than adults to

DOI: 10.1002/acp

suggestion and other social influences' was reliable enough to be presented by experts in court (Kassin, Tubb, Hosch, & Memon, 2001). Similar to previous survey respondents, experts and laypeople in our study indicated that preschoolers are less accurate when responding to misleading questions than older children and adults.

Our study improves upon extant survey research investigating individuals' perceptions of age-related differences in suggestibility in several ways. First, researchers designed earlier surveys to assess respondents' attitudes and opinions towards a wide variety of factors (e.g. witnesses' testimonial competence or the reliability of expert evidence). By design, those surveys only contained a limited number of items (one or two) that specifically addressed suggestibility. Second, two of the five surveys focused exclusively on the opinions of legal professionals and law enforcement personnel; they did not include an expert psychologist sample and only one of the surveys included a juror sample. Learning how attorneys, judges, police officers and other select groups evaluate the age-related differences in suggestibility is informative; however, only data involving expert and juror knowledge of witness suggestibility can speak to the potential helpfulness of expert testimony required by FRE Rule 702.

This is not to say that past research has neglected expert and juror knowledge of witness suggestibility altogether. At least two surveys have examined expert and juror knowledge of topics related to child sexual abuse, including children's cognitive capabilities and the reliability of their reports (Kovera & Borgida, 1997; Morison & Greene, 1992). Those studies revealed that juror and expert knowledge of child sexual abuse-related factors were sometimes at odds. At least 80% of the experts in these studies disagreed with the statement that it is easy to manipulate children into giving false reports. Jurors, in contrast, tended to agree with that statement. Based on their survey results, both research teams independently concluded that expert testimony addressing deficits in jurors' knowledge could be beneficial in child sexual abuse cases. Yet as with many of the other studies we have described, those surveys did not focus exclusively on witness suggestibility, but instead examined the reliability of children's reports within the broader context of child sexual abuse. Moreover, the restricted age range and high stress levels typically associated with child sexual abuse crimes may limit the extent to which those findings characterize expert and juror knowledge of witness suggestibility more generally.

Two additional points should be addressed before leaving our discussion of age-related differences in witness suggestibility. First, our data raise the question of why experts' effects size estimates did not mirror Ceci and Bruck's (1993) conclusion that preschoolers are more suggestible than older children and adults, who are equally suggestible. One possibility is that many years have passed since Ceci and Bruck's pioneering work and a great deal of research on age-related differences in suggestibility has occurred in the interim. Perhaps our experts' effect size estimates reflect suggestibility research that was not published or available at the time Ceci and Bruck conducted their qualitative review. It is also possible that experts responding to our survey had more or less knowledge of the developmental literature on witness suggestibility than Ceci and Bruck, interpreted that literature differently with respect to age differences in suggestibility, or varied in their conceptions of how suggestibility is defined.[4]

Second, although ~~it is true~~ both experts and laypeople recognized that preschoolers are more suggestible than older children and adults, experts' effect size estimates were nearly twice as large for each age comparison than those of laypeople. In other words, laypeople

[4]We would like to thank Robert Belli for suggesting this potential explanation.

DOI: 10.1002/acp

underestimated the size of suggestibility differences between preschoolers versus older children and adults compared to experts.

Moderators of witness suggestibility

As hypothesized, laypeople in our study lacked a sufficient understanding of many variables known to moderate the effects of misleading information on witness accuracy. Consistent with the empirical literature on witness suggestibility, experts reported that witnesses were more resistant to misinformation concerning central event details than to misinformation about peripheral event details (Heath & Erickson, 1998; Warren & Lane, 1995; Wright & Stroud, 1998). Experts indicated that witnesses who actively participated in the to-be-remembered event were more accurate when questioned in a misleading manner than witnesses who passively observed the event (Rudy & Goodman, 1991; Tobey & Goodman, 1992). Experts also estimated that a misleading suggestion from a highly prestigious source was more damaging to witness accuracy than a misleading suggestion from a source of equal or lesser status (Ceci et al., 1987; Dodd & Bradshaw, 1980; Lampinen & Smith, 1995; Smith & Ellsworth, 1987). In stark contrast, jurors and college students failed to report any such differences.

Experts and laypeople reached higher levels of agreement when estimating how two factors interact with misleading information to affect witness accuracy: a warning and the passage of time. Experts, jurors and college students agreed that warned witnesses were somewhat less suggestible than nonwarned witnesses. Similarly, all three respondent groups indicated that witness accuracy decreases when either the delay between the event and exposure to misleading information or the delay between the event and the final memory test increases. Respondents' estimates mirrored previous research findings that warnings (Christiaansen & Ochalek, 1983; Greene et al., 1982; Saywitz & Moan-Hardie, 1994; Warren et al., 1991) and shorter retention intervals (Belli et al., 1992; Loftus et al., 1978; Saywitz et al., 1991; Tucker et al., 1990) attenuate the negative effects of misinformation.

Why were jurors knowledgeable about some factors that moderate the effects of misleading information on witness accuracy and not others? Although we cannot know for sure, the effects of a warning and the passage of time seem to fall in line with intuition or 'common sense' to a greater degree than those of more highly sophisticated and nuanced suggestibility variables such as event detail centrality or participation level. That is, understanding that warnings serve a preventative function and that memory fades over time draws on more general knowledge and experience, whereas realizing that the centrality of event details or a witness's level of participation can influence accuracy requires more specialized knowledge that lay respondents simply did not possess.

We admit, however, this explanation is less satisfying when applied to the deficits observed in jurors' understanding of source prestige/authority effects on witness accuracy. The fact that a high authority source exerts a greater influence on witness suggestibility than a low authority source would appear to make intuitive sense given children's general tendency to defer to adults and authority figures. Nevertheless, jurors did not recognize the influence of this variable on witness accuracy and experts did. Interestingly, however, jurors' misunderstanding of source prestige/authority effects is consistent with the fundamental attribution error (Jones & Harris, 1967; Ross, 1977) or people's tendency to routinely underestimate situational influences on others' behaviour in favour of dispositional influences.

DOI: 10.1002/acp

Familiarity, appropriateness and helpfulness of expert testimony

Results supported our hypotheses that laypeople would rate themselves as being less familiar with witness suggestibility research compared to experts and that all respondents would view expert testimony as appropriate and beneficial to jurors. Differences in respondents' familiarity ratings were quite large, with experts and laypeople falling at opposite ends of the scale ('very familiar' and 'not at all familiar'). Differences on the appropriateness and helpfulness measures were smaller, indicating a shared belief that expert testimony on witness suggestibility is appropriate and helpful.

Compared to the other two respondent groups, jurors rated themselves as being the least familiar with suggestibility research and provided the lowest ratings (although still positive) regarding the appropriateness and helpfulness of expert testimony on that topic. On the one hand, these results are surprising: it would seem that if jurors perceived themselves as being the least familiar with suggestibility research that they would find expert testimony on this topic most beneficial. This was not the case. On the other hand, perhaps these results are not surprising. Because jurors perceived themselves to be unfamiliar with the witness suggestibility literature, they may have found it difficult to rate the appropriateness and helpfulness of expert testimony on this topic and gravitated toward the midpoints of the scales. It is also possible that jurors are more skeptical about expert testimony than others based on their perception that experts are 'hired guns' (Cooper & Neuhaus, 2000). College students may be less susceptible to this belief due to their increased classroom exposure to the witness suggestibility research that informs expert testimony.

Implications for the potential helpfulness of expert testimony on witness suggestibility

Can expert testimony on witness suggestibility tell jurors something they don't already know? Our findings provide preliminary evidence that the answer to this question is 'yes'. Compared to experts, laypeople in our survey underestimated the size of suggestibility differences between preschoolers versus older children and adults. Both jurors and college students lacked knowledge about how event detail centrality, level of witness participation, and source prestige can moderate the effects of misleading information on witness accuracy. Finally, laypeople themselves reported being largely unfamiliar with witness suggestibility research and that expert testimony on that topic would be helpful.

Expert testimony aimed at correcting jurors' tendency to underestimate the size of age-related suggestibility should be helpful in cases involving the relative suggestibility of witnesses from different age groups (e.g. a case involving conflicting testimony from a younger versus older child who were both interviewed in a suggestive manner). Also, expert testimony could assist jurors by informing them that answering questions about peripheral event details, passively observing the event, and being questioned by a prestigious source can increase witness suggestibility. Equally helpful, expert testimony could describe that answering questions about central event details, actively participating in the event, and being questioned by a source of equal or lesser status can decrease witness suggestibility in response to misleading questions.

Despite the lay knowledge deficits that emerged, jurors and college students appeared to adequately understand general age-related trends in suggestibility, the positive effects of a warning prior to misinformation exposure on witness memory, and the negative influence

DOI: 10.1002/acp

of longer delays between the event/misinformation and the event/final memory test on witness accuracy. As a result, expert testimony on these issues would not appear to be helpful in light of FRE Rule 702; consensus between expert and lay knowledge of these issues already exists.

LIMITATIONS

We must acknowledge certain limitations when considering whether our findings generalize to actual cases involving expert testimony on witness suggestibility. The first two limitations concern the representativeness of our expert sample. Although we modeled the inclusion criteria for experts after FRE Rule 702, which states that a witness must be qualified as an expert on the basis of their 'skill, knowledge, experience, training or education', it is possible that we neglected to contact some qualified experts. The primary means for compiling the names and academic affiliations of researchers to be included in the expert sample was to reference the PsychLit and Dissertation Abstracts research databases. As a result, some psychologists who possess the attributes necessary to be qualified as experts on witness suggestibility were not included in the sample simply because they have not published in the area of witness suggestibility (e.g. nonacademic clinicians who have extensive experience interviewing children in clinical settings).[5] Moreover, although we used different keywords to search the research databases, we may have overlooked the names of researchers who in fact have authored witness suggestibility publications. If sampling errors such as these occurred, they were infrequent and nondeliberate. We took great care in assembling the expert sample and used multiple methods to ensure that we included as many qualified experts as possible.

Second, despite diligent attempts to maximize the response rate for this survey by following Dillman's (1978; Salant & Dillman, 1994) Total Design Method, approximately 20% of experts surveyed failed to respond in any way. Moreover, just over half (56%) of the experts we originally contacted returned surveys that were entirely completed. Although we were able to test for differences between experts who completed the entire survey and those who returned incomplete surveys, it was impossible to test for differences between experts who responded to this survey and those who did not respond in any way. Psychologists who failed to respond to this survey may possess knowledge and attitudes that are quite different from the psychologists who participated in our study. Despite these concerns, it is reassuring that the percentage of experts responding to this survey is comparable to those observed in other similar published surveys (e.g. Kassin et al., 1989; Morison & Greene, 1992).

Practical considerations also limited our survey in several ways. First, the survey contained a finite set of factors that psychologists have studied for their potential moderating influence on the relationship between misleading information and suggestibility. Clearly other moderators exist that were not included in the survey. We did not ask respondents to estimate the effects of misinformation on witnesses who were misled once versus repeatedly, who were questioned alone versus in the presence of others, who received plausible versus implausible suggestions, who received written versus oral

[5]One reviewer argued that if these professionals were missing or underrepresented in our sample, our results may actually *underestimate* differences in expert and juror knowledge of witness suggestibility. For example, clinicians routinely testify with respect to whether a forensic interview of a child conforms to readily accepted guidelines. Because our survey did not measure respondents' knowledge of such guidelines, the data do not speak to whether experts and lay people are equally knowledgeable about this information.

Appl. Cognit. Psychol. 20: 1–22 (2006)
DOI: 10.1002/acp

misinformation, or who experienced longer or shorter pre- and post-misinformation delays. Also this study did not address the extent to which experts and laypeople believe that the type of final memory test used (e.g. the original memory test used by Loftus et al., 1978 versus the modified recognition test used by McCloskey & Zaragoza, 1985 versus the source monitoring test used by Lindsay & Johnson, 1989) moderates the presence of a misinformation effect. We opted for a shorter, more focused survey as recommended by Dillman (1978) to help obtain a higher response rate. Even the most thorough and exhaustive survey is useless if no one completes it.

Second, the survey did not allow respondents to take into account how higher-order interactions among variables may influence the misinformation/accuracy relation. Although our extensive review of the suggestibility literature revealed that studies investigating the higher-order interactions among variables are uncommon, it is likely that some of the variables we examined could interact and uniquely affect witness suggestibility. The response format of our survey did not capture these higher-order interactions.

Third, laypeople's hypothesized lack of expertise regarding the suggestibility paradigm required that we include in our survey examples of how researchers have operationalized certain variables in their studies. Doing so may have inadvertently introduced a confound into at least one of the witnessing condition item pairs: level of participation. In other words, the two witnessing condition items varied not only on participation level (actively participate versus passively observe), but also on the nature of the example provided (playing a game with researcher versus watching a crime simulation). Although these were actual examples taken from the witness suggestibility research literature, one reviewer noted that this confound may have attenuated the effects of witnessing conditions for laypeople compared to experts because the experts are more familiar with the ethical constraints of suggestibility research (i.e. not using to-be-remembered events that physically or psychologically harm participants) than jurors.

We cannot entirely dismiss this possibility, but would like to note this limitation goes to the interpretation of a difference rather than to the determination of whether a difference existed. Laypeople and experts differed in their understanding of participation level and witness suggestibility irrespective of whether that difference was due to the witnessing conditions or examples provided. Because this difference exists, expert testimony addressing the effects of participation level on witness suggestibility (which presumably would include some explanation of how those variables were operationalized in the research) should benefit jurors from a FRE Rule 702 standpoint.

Two final limitations should be kept in mind when considering the implications of our results. First, many of the differences observed (although statistically significant) were small in size with partial η^2 values < 0.10. Second, our data represent a snapshot in time of expert and lay beliefs about witness suggestibility. As future research continues to improve our understanding of witness suggestibility, our data (particularly those measuring experts' beliefs) will have to be updated to ensure that they truly reflect respondents' knowledge and the current state of social scientific literature.

CONCLUSION

Certain information about witness suggestibility exceeds jurors' common knowledge and understanding. One obvious way to address jurors' deficits is for courts to allow experts to testify in cases when the suggestibility of a witness—young or old—is at issue. Under

DOI: 10.1002/acp

FRE Rule 702, experts could inform jurors about certain factors that may increase or decrease the effects of misinformation on witness suggestibility. Such information *should* improve jurors' decisions in witness suggestibility cases; however, only future research will reveal if this is indeed the case. In addition, expert testimony may be excluded if its probative value is substantially outweighed by its potential to unfairly prejudice, confuse or mislead the jury (Rule 403). Thus, expert testimony addressing jurors' knowledge deficits about witness suggestibility may not be admitted in actual cases if it prejudices, confuses, or misleads triers of fact. Additional studies should scrutinize whether expert testimony on witness suggestibility detrimentally affects juror decision-making in this or any other way.

ACKNOWLEDGEMENTS

This research was conducted in partial fulfilment of the requirements for Bradley D. McAuliff's master's thesis at Florida International University, under the direction of Margaret Bull Kovera. Parts of this research were presented at the biennial meeting of the American Psychology-Law Society in Redondo Beach, CA (March, 1998) and the 106th convention of the American Psychological Association, San Francisco, CA (August, 1998). This research was supported by a grant-in-aid from the American Psychology-Law Society. The authors thank the jury administration staff and personnel at Broward County Courthouse, Seventeenth Judicial District, Fort Lauderdale, Florida (Honourable Robert L. Andrews, Dolly Gibson, Norman Houghtaling, Pat Todaro, Lisa Muggeo, Audrey Edwards, Marvin Edelstein and Allison Mitchel) for granting access to the jury pool. The authors also thank thesis committee members Brian L. Cutler and Janat F. Parker for their helpful suggestions and Robert Schopp for his invaluable insight regarding the vagaries of evidence law.

REFERENCES

Barlow v. State, 270 Ga. 54 (1998).

Belli, R. F., Windschitl, P. D., McCarthy, T. T., & Winfrey, S. E. (1992). Detecting memory impairment with a modified test procedure: Manipulating retention interval with centrally presented event items. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 356–367.

Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior, 23*, 75–91.

Brigham, J. C., & Spier, S. A. (1992). Opinions held by professionals who work with child witnesses. In H. Dent & R. Flin (Eds.), *Children as witnesses* (pp. 93–111). Chichester, England: Wiley.

Bruck[Q3], M., & Ceci, S. J. (1995). Amicus brief for the case of *State of New Jersey v. ~~Margaret Kelly~~ Michaels* presented by Committee of Concerned Social Scientists. *Psychology, Public Policy, and Law, 1*, 272–322.

Bruck, M., & Melynk, L. (2005). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology, 18*, 947–996.

Cashmore, J., & Bussey, K. (1996). Judicial perceptions of child witness competence. *Law and Human Behavior, 20*, 313–334.

Ceci, S. J., & Bruck, M. B. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403–439.

Ceci, S. J., Ross, D. F., & Toglia, M. P. (1987). Suggestibility of children's memory: Psycholegal implications. *Journal of Experimental Psychology: General, 116*, 38–49.

Christiaansen, R. E., & Ochalek, K. (1983). Editing misleading information from memory: Evidence for the coexistence of original and postevent information. *Memory and Cognition, 11*, 467–475.

Christiaansen, R. E., Sweeney, J. D., & Ochalek, K. (1983). Influencing eyewitness descriptions. *Law and Human Behavior, 7*, 59–65.

DOI: 10.1002/acp

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
Cohen, R. L., & Harnick, M. A. (1980). The susceptibility of child witnesses to suggestion. *Law and Human Behavior, 4*, 201–210.
Cooper, J., & Neuhaus, I. M. (2000). The 'hired gun' effect: Assessing the effect of pay, frequency of testifying, and credentials on the perception of expert testimony. *Law and Human Behavior, 24*, 149–171.
Coxon, P., & Valentine, T. (1997). The effects of the age of eyewitnesses on the accuracy and suggestibility of their testimony. *Applied Cognitive Psychology, 11*, 415–430.
Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S.Ct. 2786 (1993).
Dillman, D. A. (1978). *Mail and telephone surveys: The total design method*. New York: Wiley.
Dodd, D. H., & Bradshaw, J. M. (1980). Leading questions and memory: Pragmatic constraints. *Journal of Verbal Learning and Verbal Behavior, 19*, 695–704.
Federal Rules of Evidence, St. Paul Minnesota: West (2001).
Frazier, P., & Borgida, E. (1988). Juror common understanding and the admissibility of rape trauma syndrome evidence in court. *Law and Human Behavior, 12*, 101–121.
Frye v. United States, 293 F. 1013 (1923).
Greene, E., Flynn, M. S., & Loftus, E. F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behavior, 21*, 207–219.
Groscup, J. L., Penrod, S. D., Studebaker, C. A., Huss, M. T., & O'Neil, K. M. (2002). The effects of *Daubert* on the admissibility of expert testimony in state and federal criminal cases. *Psychology, Public Policy, and Law, 8*, 339–372.
Heath, W. P., & Erickson, J. R. (1998). Memory for central and peripheral actions and props after varied post-event presentation. *Legal and Criminological Psychology, 3*, 321–346.
Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology, 3*, 2–24.
Kassin, S. M., Ellsworth, P. C., & Smith, V. L. (1989). The 'general acceptance' of psychological research on eyewitness testimony: A survey of the experts. *American Psychologist, 44*, 1089–1098.
Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the 'general acceptance' of eyewitness testimony research: A new survey of the experts. *American Psychologist, 56*, 405–416.
Kovera, M. B., & Borgida, E. (1997). Expert testimony in child sexual abuse trials: The admissibility of psychological science. *Applied Cognitive Psychology, 11*, S105–S129.
Lampinen, J. M., & Smith, V. L. (1995). The incredible (and sometimes incredulous) child witness: Child eyewitnesses' sensitivity to source credibility cues. *Journal of Applied Psychology, 80*, 621–627.
Leippe, M. R., Brigham, J. C., Cousins, C., & Romanczyk, A. (1989). The opinions and practices of criminal attorneys regarding child eyewitnesses: A survey. In S. J. Ceci, D. F. Ross, & M. P. Toglia (Eds.), *Perspectives on children's testimony* (pp. 100–130). New York: Springer-Verlag.
Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory and Cognition, 17*, 349–358.
Loftus, E. F., & Greene, E. (1980). Warning: Even memory for faces may be contagious. *Law and Human Behavior, 4*, 323–334.
Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 4*, 19–31.
McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General, 114*, 1–16.
Morison, S., & Greene, E. (1992). Juror and expert knowledge of child sexual abuse. *Child Abuse and Neglect, 16*, 595–613.
People v. Behrooz Kanani, 272 A.D.2d 186 (1st Dept 2000).
People v. Timothy Johnston, 273 A.D.2d 514 (3rd Dept 2000).
Quas, J. A., Thompson, W. C., & Clarke-Stewart, K. A. (2005). Do jurors 'know' what isn't so about child witnesses? *Law and Human Behavior, 29*, 425–456.
Ross, D. F., Dunning, D., Toglia, M. P., & Ceci, S. J. (1990). The child in the eyes of the jury: Assessing mock jurors' perceptions of the child witness. *Law and Human Behavior, 14*, 5–23.
Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10). New York: Academic Press.

Appl. Cognit. Psychol. 20: 1–22 (2006)
DOI: 10.1002/acp

Rudy, L., & Goodman, G. S. (1991). Effects of participation on children's reports: Implications for children's testimony. *Developmental Psychology, 4*, 527–538.
Salant, P. A., & Dillman, D. A. (1994). *How to conduct your own survey.* New York: Wiley.
Saywitz, K. J., Goodman, G. S., Nicholas, E., & Moan, S. F. (1991). Children's memories of a physical examination involving genital touch: Implications for reports of child sexual abuse. *Journal of Consulting and Clinical Psychology, 59*, 682–691.
Saywitz, K. J., & Moan-Hardie, S. (1994). Reducing the potential for distortion of childhood memories. *Consciousness and Cognition, 3*, 408–425.
Smith, V. L., & Ellsworth, P. C. (1987). The social psychology of eyewitness accuracy: Misleading questions and communicator expertise. *Journal of Applied Psychology, 72*, 294–300.
State of New Jersey v. Margaret Kelly Michaels, 136 N.J. 299, 642 A.2d 1372 (N.J., 1994).
State of Ohio v. Gersin, 76 Ohio St.3d 491 (1996).
Swim, J. K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *Journal of Personality and Social Psychology, 66*, 21–36.
Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson Education.
Tobey, A., & Goodman, G. S. (1992). Children's eyewitness memory: Effects of participation and forensic context. *Child Abuse and Neglect, 16*, 779–796.
Tucker, A., Mertin, P., & Luszcz, M. (1990). The effect of a repeated interview on young children's eyewitness memory. *Australian and New Zealand Journal of Criminology, 23*, 117–124.
Warren, A. R., & Lane, P. (1995). The effects of timing and type of questioning on eyewitness accuracy and suggestibility. In M. Zaragoza, et al. (Eds.), *Memory and testimony in the child witness* (pp. 44–60). Thousand Oaks, CA: Sage.
Warren, A., Hulse-Trotter, K., & Tubbs, E. C. (1991). Inducing resistance to suggestibility in children. *Law and Human Behavior, 15*, 273–285.
Wright v. State, 233 Ga. App. 358 (1998).
Wright, D. B., & Stroud, J. T. (1998). Memory quality and misinformation for peripheral and central objects. *Legal and Criminological Psychology, 3*, 273–286.
Yuille, J. C. (1989). Expert evidence by psychologists: Sometimes problematic and often premature. *Behavioral Sciences and the Law, 7*, 181–196.

APPENDIX A: WITNESS SUGGESTIBILITY SURVEY (EXPERT VERSION)

On the following pages you will find a series of statements describing different variables that have been studied for their potential influence on the suggestibility of witnesses. Please read these statements carefully and indicate how large or small you think the difference is between the accuracy of misled and nonmisled witnesses in each of the following situations. For the purposes of this survey, please understand that the term 'misled' refers to those individuals who received some form of misleading or incorrect information about an event after witnessing that event. The term 'nonmisled' refers to those individuals who witnessed the same event as the misled individuals, yet received no misleading information about that event.

You will be estimating the size of the difference in the accuracy of misled and nonmisled witnesses for three primary age groups: preschoolers (0 to 5 years old), elementary/middle school children (6 to 13 years old), and high school students/adults (14 years and older). Please use the guidelines below to help indicate your estimated size of the memory difference between misled and nonmisled witnesses:

I believe that the difference in the accuracy of misled and nonmisled witnesses is . . .

Misled more accurate than nonmisled				Misled less accurate than nonmisled		
Large	Moderate	Small	None	Small	Moderate	Large
−0.80	−0.50	−0.20	0	0.20	0.50	0.80

Appl. Cognit. Psychol. 20: 1–22 (2006)
DOI: 10.1002/acp

Please keep in mind that negative numbers indicate that misled individuals are more accurate than nonmisled individuals, whereas positive numbers indicate that nonmisled witnesses are more accurate than misled witnesses.

Realize you are not restricted to using only the numbers listed above. These numbers were provided for the sole purpose of helping you gauge how different the accuracy of misled and nonmisled witnesses is. If you believe that misled witnesses are somewhat more accurate than are nonmisled witnesses, you might estimate the size of this difference to be −0.25. If you believe that nonmisled witnesses are much more accurate than are misled witnesses, you might estimate the size of this difference to be 2.00. Depending on the presence or absence of certain variables, differences of all sizes have been found in studies of witnesses' susceptibility to misinformation.

You may or may not see fit to adjust the size of the differences in the accuracy of witness memory according to the age of the witness. For example, you may believe that misled witnesses, irrespective of their age, have less accurate memories for an event than do nonmisled witnesses, you might mark down 0.40 for all three age groups. If you believe that the difference in the accuracy of misled and nonmisled witnesses is larger for high school students/adults than for elementary/middle school children, you might answer 1.00 for the high school students/adults and 0.80 for the elementary/middle school children.

APPENDIX B: QUESTIONS USED TO ASSESS CONSTRUCT VALIDITY OF WITNESS SUGGESTIBILITY SCALE

Assume the following hypothetical statement is true: We know that in a certain set of circumstances, misled witnesses are less accurate than are nonmisled witnesses. When estimating the size of this difference using the scale provided, would your estimate values be positive or negative?

Assuming the same information above is true, if the difference in the accuracy of misled and nonmisled witnesses is medium for preschoolers (0–5 years old), large for elementary/middle school children (6–13 years old), and small for high school students/adults (14 years and older), what would your estimate values be using the scale provided?

_____ for preschoolers (0–5 years old)
_____ for elementary/middle school children (6–13 years old)
_____ for high school students/adults (14 years and older)

Note. These questions were included in 60% of the college student surveys. Responses were combined and coded as either correct or incorrect. Any question that contained partially incorrect information was coded as an incorrect response.

DOI: 10.1002/acp

Author Query Form (ACP/1301)

Special instructions: Please list responses in an e-mail. Alternatively please write responses to queries directly on the Galley proofs and fax back or scan and attach in an e-mail.

Q1: Author: Please check the insertion of year in the reference "Federal Rules of Evidence" as per the list. FINE

Q2: Author: Please check the year of the "Tabachnick" reference. 2007 IS CORRECT FOR 5TH EDITION

Q3: Author: Reference not cited in the text. Please check.

ADDED ON P. 6