Child Eyewitnesses: Seeing Is Believing

C. A. Elizabeth Luus
University of Victoria

Gary L. Wells
Iowa State University

John W. Turtle
York University

Two studies examined perceptions of child and adult eyewitnesses’ credibility. In Study 1, college students evaluated transcribed testimonies of 8-year-old and adult witnesses to a videotaped staged crime. Half were misinformed about the witness’s age (i.e., either believing a child’s testimony was provided by an adult or vice versa). Neither actual age nor ostensible age affected participants’ evaluations. In Study 2, adults (N = 85) viewed videotaped testimonies of 8-year-old, 12-year-old, and adult eyewitnesses. Half viewed 1 witness’s testimony and then evaluated his or her credibility. The others viewed only a still frame of 1 witness, then imagined the testimony the witness had provided, and finally evaluated him or her credibility without having actually heard the testimony. Young children were judged more favorably when their entire testimony, rather than a still frame, was viewed. This was not true for the older eyewitnesses.

What cues do jurors use to judge the credibility of child eyewitnesses? Are jurors’ evaluations of children’s testimony biased by negative stereotypes of children’s eyewitness capabilities? This article addresses these questions by comparing participant jurors’ evaluations of their imagined versus the actual eyewitness capabilities of samples of child and adult eyewitnesses.

The legal community has historically regarded the testimony of children with distrust, believing children are too suggestible or prone to incorporate misleading suggestions into their memory reports (Whipple, 1911). Surveys of adults’ beliefs about child eyewitnesses have suggested that the general population shares this pessimistic view of children’s eyewitness capabilities (e.g., Leippe & Romanczyk, 1987; Ross, Dunning, Toglia, & Ceci, 1991; Yarmey & Jones, 1983). For example, respondents to Ross et al.’s survey reported that they believed a child eyewitness would be less accurate and more suggestible than an adult eyewitness. They also reported that they would assign less weight to testimony if it was provided by a child rather than an adult.

Experimental research has confirmed that child eyewitnesses are more suggestible than their adult counterparts (e.g., Ceci, Ross, & Toglia, 1987; Cohen & Harnick, 1980; King & Yuille, 1987; Wells, Turtle, & Luus, 1989). For example, Wells et al. had samples of children (8- and 12-year-olds) and adults watch a simulated-crime videotape and then respond to direct- and cross-examination questioning about the crime. Cross-examination differs from direct examination in several ways. Direct examination is conducted by the attorney who calls the witness to the stand. It is generally a friendly exchange in which the examiner tries to make the witness look credible while observing the rule against leading the witness. Cross-examination is conducted by the opposing counsel, whose interest is in undermining the perceived credibility of the eyewitness. The cross-examiner is allowed to use leading questions.

Wells et al.’s (1989) direct examination consisted of a series of straightforward, nonleading questions; the cross-examination included a number of leading items. The children were as accurate as the adults in responding to direct-examination questioning. However, the children were significantly less accurate than were the adults under cross-examination. These findings are consistent with other studies that have found that the accuracy of child and adult eyewitnesses is about the same under nonlead-
ing, free-recall formats (e.g., Marin, Holmes, Guth, & Kovac, 1979; Saywitz, 1987), but child eyewitnesses are less accurate than adults in response to leading questions (e.g., Ceci et al., 1987; Cohen & Harnick, 1980; King & Yuille, 1987; Leippe, Manion, & Romanczyk, 1992).

Do adult observers perceive these differences in the accuracy of cross-examination testimony provided by children versus adults? That is, are the children providing discernible cues to the weakness of the cross-examination testimonies? Empirical investigations of this question have yielded mixed results. Some studies have found that child eyewitnesses are perceived as less credible than adults (e.g., Goodman, Golding, & Haith, 1984; Goodman, Golding, Helgeson, Haith, & Michelli, 1987; Leippe et al., 1992; Leippe & Romanczyk, 1987, Study 2 of 1989). Others have found that child witnesses are perceived as more credible than adults (e.g., Study 4 of Leippe & Romanczyk, 1989; Ross, Miller, & Moran, 1987). Still others have found no age-related differences in the perceived credibility of child and adult eyewitnesses (Wells et al., 1989). Wells et al. found that for both adult and child eyewitnesses, participant jurors tended to believe those eyewitnesses whom they perceived as confident. Thus, 8-year-old witnesses perceived as highly confident were rated as more credible than many adult witnesses who apparently lacked confidence.

How can these contradictory findings be reconciled? We see the inconsistencies as linked to several discrepancies in the methods used in the studies. First, the studies vary in terms of the comprehensiveness of the testimony participant jurors were asked to evaluate. Some studies assessed jurors' stereotypes of children's eyewitness capabilities by providing participant jurors with (a) a description of a case with only a statement of the witness's age, (b) a written trial transcript (and manipulated participant jurors' assumptions about the age of the witness), or (c) a videotape of a child or an adult actor reciting a testimony script. Other studies (Leippe et al., 1992; Wells et al., 1989) have gone beyond the question of adults' stereotypes of children's eyewitness capabilities by presenting participant jurors with samples of child and adult eyewitnesses who actually witnessed an event and then responded to questions concerning their memory for the event.

The stereotype studies have obtained dramatically different results from one another. For example, whereas Goodman et al. (1987) found their videotaped child eyewitness to be perceived as less credible than their adult eyewitness, Ross et al. (1987) found their child eyewitness to be perceived as more credible than their adult eyewitness. To explain these inconsistent findings, we note that previous research has revealed profound differences across eyewitnesses in their perceived credibility (e.g., Wells, Lindsay, & Ferguson, 1979; Wells et al., 1989). The idea that there is variance in the perceived credibility of eyewitnesses within ages suggests the possibility that some child eyewitnesses might be perceived as more credible than some adult eyewitnesses and vice versa (as found by Wells et al., 1989). It then follows that a single example of a child's or an adult's testimony cannot be considered representative of that age group's testimony in general. Thus, the specific child or adult actor who delivered the testimony could have made all the difference in the varying results obtained by the stereotype studies.

Provision of reports from children who are genuinely trying to recall a witnessed event from memory using their own words is, in our opinion, an important feature of any research concerned with perceptions of children's eyewitness credibility. To date, only two studies (Leippe et al., 1992; Wells et al., 1989) have included this feature in their research. These studies varied in terms of the stimulus materials used (Wells et al. had participants view a simulated-crime film, whereas Leippe et al. tested participants' memory for a live interaction involving a bogus skin-sensitivity test), the questions included in the memory test, and the age of participants tested (Leippe et al. included 5–6-year-olds in their sample; Wells et al.'s youngest participants were 8 years old). Although the two studies were not entirely consistent in their overall findings, they both found a tremendous amount of variation in perceived credibility within their samples of child and adult participant witnesses. "On average, adults may appear more credible, but some significant number of children may look just as credible—or moreso—than some adults" (Leippe et al., 1992, p. 190). This was certainly the case with the children who served as witnesses in Wells et al.'s study.

Wells et al. (1989) found that eyewitness age was unrelated to any of the credibility assessments made by participant jurors. Participant jurors were sensitive to the main effect difference between direct-examination and cross-examination performance (appropriately estimating lower accuracy for cross-examination than for direct examination) but were not sensitive to the interaction between type of examination (direct vs. cross) and eyewitness age. For both adult and child eyewitnesses, participant jurors tended to believe those eyewitnesses whom they perceived as confident. Thus, interestingly, 8-year-old witnesses who were perceived as highly confident were rated as more credible than those adult witnesses who displayed lower levels of confidence.

These results lend credence to Leippe and Romanczyk's (1989) conclusion that "adults' negative preconceptions about children's memory will not dispose them to reject a child's memory message if the message's quality is sufficiently 'mature' to belie the stereotype" (p. 127). These researchers presented participant jurors
with a description of a case involving a child or adult witness or they had participant jurors read a transcript of the child or adult witness's testimony. In general, their results showed that participant jurors tended to show a bias against the child’s testimony in the description cases in which they merely knew the child’s age but a bias favoring the child’s testimony when they read the testimony transcript. This finding is consistent with the notion of augmentation in attribution theory (Kelley, 1972) in which a behavior that occurs in spite of some inhibitory factor (rather than because of a facilitory factor) is perceived to have an especially strong underlying cause. In this case, a child’s young age may have been perceived as an inhibitory factor, and hence, perceptions of the child’s memory may have been augmented relative to adults’ memory.

Leippe and Romanczyk’s (1989) findings have suggested that it is possible that jurors may imagine children to be incapable of delivering accurate, convincing testimony. However, they may come to reconsider these preconceived negative views of child eyewitnesses upon actually observing a child testify. Our research builds on Leippe and Romanczyk’s findings in addressing this possibility.

In our first experiment, we presented participant jurors with testimony transcripts from a sample of 8-year-old eyewitnesses and a sample of adults to see if they could discern differences in the verbal portions of the testimony. By using transcripts rather than videotapes, we were able to manipulate the actual age of the witnesses independently of the ostensible age of the witnesses. Thus, some participant jurors evaluated actual 8-year-olds’ testimonies, thinking that the witnesses were 25 years old, whereas others evaluated actual 25-year-olds’ testimonies, thinking that the witnesses were 8 years old, and so on.

In our second experiment, a set of 42 testimony videotapes of witnesses who took part in a preliminary study (described below) served as stimulus materials. Participant jurors either viewed an 8-year-old, a 12-year-old, or an adult witness and evaluated his or her credibility or viewed only a still frame of one witness and evaluated his or her credibility without hearing the actual testimony. In the latter condition, participants were instructed to imagine the testimony the witness would provide. This procedure allowed us to directly compare participant jurors’ perceptions of real versus imagined eyewitness capabilities of child and adult eyewitnesses.

Preliminary Study: Creating Stimulus Materials

We had 8-year-olds, 12-year-olds, and adults \( n = 14 \) per age group individually witness a filmed abduction of a child from a playground. Participants were not forewarned of an upcoming memory test. We were sensitive to the fact that this event might evoke fear among the children. The video we created depicted a nonviolent scene in which a man approached a young boy, talked to him in a friendly manner, then offered him some money and coaxed him to go for a drive in his car. The boy hesitantly but willingly accompanied the man.

The parents of all child participants were fully informed of the content of the video through a letter that was sent home with all students in participating classrooms. Parents indicated their willingness to allow their child to participate by signing and returning an attached consent form. The principals and the parent–teacher associations of the participating schools viewed the crime video before granting their approval of the study. In debriefing the young children, we discussed playground safety and reviewed some measures the children might take to remember details of events they suspect may be dangerous to themselves or others (e.g., try to write down the license number of suspicious cars).

The next day, each eyewitness responded to direct- and cross-examination questioning while being videotaped. The direct examination consisted of straightforward questions (e.g., “How many children were at the playground?”). The cross-examination questions were not of this straightforward nature but rather constituted an attempt to undermine the credibility of the witness, to elicit inconsistent testimony, or both. For example, one of the cross-examination questions was “You claimed before that the playground was fairly crowded, is that correct?” In fact, only three people were at the playground (and nearly every witness had previously made this assertion). Both the direct- and cross-examination questions were scripted in a way that allowed the questioners to follow their examination format regardless of how the witness responded. These testimony videotapes served as stimulus materials in our research.

The children were sampled from four different public schools located in different areas of a medium-sized city. We chose schools that varied in terms of the affluence of the surrounding neighborhoods to sample a diversity of socioeconomic backgrounds. Furthermore, we avoided selecting children who were not representative of their age group in that we did not sample the children only from advanced-achievement or only from lower achievement school programs.

Our decision to study 8- and 12-year-olds in comparison with adults was based primarily on current knowledge of children’s understanding of the legal system and their cognitive capabilities. Studies of reality monitoring (e.g., Johnson & Foley, 1984; Johnson & Raye, 1981) have suggested that children develop the ability to distinguish fantasy from reality between 6 and 8 years of age. By age 8, children have acquired an understanding of
rights (Melton, 1980) and a view of both their general self-worth and their competence in different domains (Harter, 1982). At this age, children also begin to comprehend concepts of court and to understand the roles played by judges, lawyers, and witnesses (Saywitz, 1989). However, the function of the jury and a sense of the legal system’s role in society are not appreciated until approximately 12 years of age (Saywitz, 1989).

All witnesses responded to 17 questions: 10 direct-examination questions and 7 cross-examination questions. Answers to these questions were scored by determining definite boundaries on the range of possible replies to each question. For example, “burgundy” and “purple” were the only acceptable descriptions for the color of the getaway car. Interjudge reliability was virtually perfect because the scoring criteria were very explicit. Univariate analyses of variance (ANOVAs; with age as the factor) on the total number of questions answered correctly indicated that accuracy did not differ significantly with age for the direct-examination responses, \( F(2, 39) = 2.23, p = .12 \) There was, however, a significant age difference in the accuracy of the cross-examination responses, \( F(2, 39) = 8.91, p = .007 \). A subsequent Newman–Keuls analysis revealed that the 8-year-old children were significantly less accurate under cross-examination (mean number of correct responses = 3.07, \( SD = 0.92 \)) than were the 12-year-old and adult witnesses \((p < .01)\). The 12-year-old and adult eyewitnesses did not differ in the accuracy of their cross-examination responses (for the 12-year-olds, the mean number of correct responses = 4.57, \( SD = 1.28 \); for the adults, the mean number of correct responses = 4.28, \( SD = 1.07 \)).

An item analysis of the cross-examination questions revealed that the young children evidenced a pattern of erring in response to the leading questions. The adolescents and adults did not show a similar pattern of question-specific errors. For example, in response to the query, “Was the man carrying his wallet in his right hand or his left hand?” 100% of the 8-year-olds replied either “left” or “right.” The man, in fact, did not have a wallet. Thus, a correct response required correcting the questioner. For both the adolescents and the adults, 30% correctly asserted that the man had not been carrying a wallet.

The large decline in the accuracy of the 8-year-olds’ responses under cross-examination is consistent with Ceci et al.’s (1987) findings regarding young children’s susceptibility to misleading information. Ceci et al. found that very young children were significantly more likely than were adolescents to incorporate misleading information into their responses when the misinformation was conveyed by an adult. However, when the misleading information was conveyed by a 7-year-old rather than an adult, the young children were significantly less likely to include it in their reports. This finding suggests that the young children’s suggestibility was in part due to their readiness to comply with an adult authority figure. Our cross-examination questioning can be considered analogous to Ceci et al.’s misleading-questions environment in that in both situations, an accurate response required disagreeing with or contradicting an adult. Thus, the reduced accuracy of our 8-year-old participants’ cross-examination testimonies likely stemmed from the social pressure inherent in that questioning environment.

To assess the credibility of the videotaped testimonies, we had participant jurors \((N = 294)\) view the testimonies and assess the credibility of each eyewitness (see Table 1 for a summary of these credibility assessments). Eyewitness age proved to be unrelated to any of the eight credibility measures. For our first experiment, we transcribed the testimonies of a subset of the 8-year-olds and a subset of the adults. Details concerning these subsets of testimonies are provided below.

### Experiment 1

**Overview**

When evaluating the credibility of a child eyewitness versus an adult eyewitness, a person may invoke different
norms (other children and other adults, respectively), which serve as anchors for judging the credibility of these witnesses. This suggestion is consistent with the idea of a norm-comparison effect (Kahneman & Miller, 1986). Consider the sentence “The small man held the large newborn baby.” This sentence is understood without difficulty because it simultaneously evokes two different norms for size. The baby is not assumed to be bigger than the man, even though the baby is described as being large and the man is described as being small, because the baby is assumed to be compared with a different norm (other babies) than the man (other men). Our first experiment allowed us to test for norm-comparison effects in judgments of eyewitness credibility by disentangling participants’ age-related expectations from any actual age-related differences in witnesses’ testimonies.

Method

Participants and design. Introductory psychology students (N = 216), participating in partial fulfillment of a course requirement, were randomly assigned to 1 of 18 conditions in a 2 (actual age of eyewitness: 8 years old vs. adult) × 3 (ostensible age of eyewitness: 8 vs. 12 vs. 25 years old) × 3 (testimony credibility: high vs. medium vs. low) between-subjects factorial design. The latter factor, testimony credibility, was a categorical classification of the transcripts based on the composite credibility scores that the raters in our preliminary study assigned to the videotaped testimonies. The purpose of this factor was to allow us to sample from the transcripts (rather than use all 42) while at the same time to perform an ANOVA test to see if the participants who were judged most and least credible within each age group by raters in our preliminary study (on the basis of evaluations of their videotaped testimonies) were also the most and least credible when participant jurors were evaluating only written transcript versions of their testimony.

Procedure. Participants were asked to read 1 of the 18 testimony transcripts provided by eyewitnesses of the filmed abduction. Three transcripts were randomly selected from Grade 3 students who had been assigned a high-, medium-, or low-credibility score by participant jurors (as described above). Three transcripts were selected from adult witnesses who were judged by participant jurors in our preliminary study to be high, moderate, or low in terms of credibility. Three versions of each testimony were created by the written assertion that the witness was 8, 12, or 25 years old. Participants were asked to respond to questions concerning (a) the number of questions answered correctly by the witness under direct examination, (b) the number of questions answered correctly by the witness under cross-examination, (c) the believability of the witness under direct examination, (d) the believability of the witness under cross-examination, (e) the confidence of the witness under direct examination, (f) the confidence of the witness under cross-examination, (g) the overall believability of the eyewitness, and (h) the overall confidence of the eyewitness. All confidence and believability assessments were made on 7-point scales ranging from 1 (not at all) to 7 (extremely).

Results and Discussion

We conducted univariate ANOVAs on each of the eight measures. A significant main effect for the credibility classification emerged on seven of the eight measures (.01 < ps < .045). The only measure that did not show significant differences for the credibility-classification variable was the measure of believability under direct examination (p < .40). Recall that the credibility-classification variable refers to the overall perceived credibility of the witnesses as evaluated by participant jurors who viewed their videotaped testimonies in our preliminary study. This result, therefore, shows that participant jurors in this experiment were able to see the same kinds of credibility differences between witnesses merely on the basis of the transcripts that the participant jurors saw in the videotaped testimony from which these transcripts were drawn.

There were no interaction effects associated with the credibility-classification variable. There also were no main effects or interaction effects associated with actual age (all ps > .4). Thus, the 8-year-olds’ transcripts were judged to be just as credible as the adults’ transcripts.

More important, there were also no main effects or interaction effects associated with ostensible age (all ps > .5). Participant jurors did not judge the witness to be less credible when described as 8 years old than when described as 12 or 25 years old. These results refute the suggestion that norm-comparison effects underlie participant jurors’ assessments of eyewitness credibility. If the stated age of the witness had prompted participant jurors to use different standards for judging testimony credibility, then we should have found an ostensible-age effect. We conclude, therefore, that a witness’s age plays a less important role in jurors’ assessments of the witness’s credibility than does the manner in which the witness answers questions under direct examination and cross-examination.

Experiment 2

Overview

The results of our first experiment support the views of Leippe and Romanczyk (1989) that there may be a stereotyped bias against the child eyewitness, but the provision of children’s actual testimony seems sufficient to overcome these preconceptions. Our second experiment addressed the question of the fit between jurors’ imagined versus actual perceptions of the credibility of child and adult eyewitnesses.

Previous research has examined either adults’ stereotypes of children’s testimony (e.g., Ross et al., 1987) or adults’ evaluations of videotaped testimony actually pro-
vided by child eyewitnesses (Leippe et al., 1992; Leippe & Romanczyk, 1989; Wells et al., 1989). The design of the current research, however, permits a direct comparison of adults' stereotypes of children's eyewitness capabilities with adults' evaluations of actual testimony provided by child versus adult eyewitnesses.

The entire set of testimony videotapes (n = 14 per age group: 8-year-olds, 12-year-olds, and adults) served as stimulus materials in this experiment. Participants were asked either to view a videotape of the testimony of one witness and to evaluate his or her credibility or to view only a still frame of one witness and to evaluate his or her credibility without hearing the actual testimony. In the latter condition, participants were instructed to imagine the testimony the witness would provide. This procedure allowed us to directly compare participant jurors' perceptions of real versus imagined eyewitness capabilities of child and adult eyewitnesses.

Method

Participants. Eighty-five undergraduate students participated in exchange for extracredit in an undergraduate psychology course.

Procedure. Participants took part in 2–4-person groups. They were seated alone in a cubicle with a television monitor and were told that they would be shown an eyewitness to a simulated-crime film. Half of the participants were told that they would be given the opportunity to view this witness's testimony concerning the crime. The other half of the participants were told that they would be shown a still frame of the witness and would be asked to imagine the testimony this witness had provided. After either viewing or imagining the witness's testimony, participants were asked to judge (a) the number of questions answered correctly by the witness under direct examination, (b) the number of questions answered correctly by the witness under cross-examination, (c) the believability of the witness under direct examination, (d) the believability of the witness under cross-examination, (e) the confidence of the witness under direct examination, (f) the confidence of the witness under cross-examination, (g) the overall believability of the eyewitness, and (h) the overall confidence of the eyewitness. All confidence and believability assessments were made on 7-point scales ranging from 1 (not at all) to 7 (extremely).

Results

The dependent variables were highly intercorrelated (see Table 2). A multivariate analysis of variance (MANOVA) on the eight dependent measures with eyewitness age and presentation mode as between-subjects variables indicated a significant main effect of presentation mode, Wilks's lamda = 0.81, F(8, 77) = 2.21, p < .05, and a significant Eyewitness Age X Presentation Mode interaction, Wilks's lamda = 0.71, F(16, 156) = 1.79, p < .05. Eyewitness age failed to produce a significant effect, p > .1. A summary of the perceived-credibility ratings showing significant effects is provided in Table 3.

Given the strong intercorrelation among the dependent measures, we computed a single composite dependent variable for analysis. (Note that we first converted the ratings to z scores to control for the different indices used, i.e., a 10-point scale for direct-examination accuracy estimates and 7-point scales for the remaining dependent measures.) An ANOVA on this composite score yielded a significant Eyewitness Age X Presentation Mode interaction, F(2, 84) = 226.25, p < .001. Eyewitness age and presentation mode did not yield significant effects (p > .1).

Post hoc t tests yielded no significant differences between participant jurors' evaluations of the still frame versus full testimony of the 12-year-olds and the adults. Participant jurors who viewed the full testimony of an 8-year-old rated the witness as more credible than did participant jurors who only imagined an 8-year-old witness's testimony, t(84) = 3.60, p < .01. These results are displayed in Figure 1.

General Discussion

The results of Experiment 2 suggest that adults' estimates of the credibility they would assign to a child eyewitness fall short of the credibility they assign upon actually viewing a child testify. We found that eyewitness age typically interacted with testimony presentation mode (i.e., whether participants viewed the full testimony or a still frame of a participant witness). Participants who evaluated the still frame of an 8-year-old and then imagined that witness's testimony assigned significantly lower testimony credibility ratings than did participants who actually viewed the 8-year-old's testimony. Participants who evaluated the still frame of a 12-year-old or an adult generally assigned comparable or slightly higher credibility ratings to them when asked to imagine rather than actually view their testimonies.

Some legal jurisdictions have started to allow child eyewitnesses to testify without being subject to cross-examination (e.g., Germany and Israel). There is an assumption in these jurisdictions that cross-examination presents special problems for children—that it will inappropriately confuse them and shatter their apparent confidence and credibility. We found no support for the contention that cross-examination serves to shatter the perceived confidence or credibility of children's oral testimony. In fact, participant jurors who had the opportunity to view the witnesses' full testimonies rated the 8-year-olds as more accurate and equally confident and believable under cross-examination as the adults. These results were obtained in spite of the fact that the participant
Table 2
Correlations Among Measures: Study 2

<table>
<thead>
<tr>
<th>Dependent measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Direct-examination accuracy</td>
<td></td>
<td></td>
<td></td>
<td>.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Direct-examination believability</td>
<td>.61</td>
<td></td>
<td></td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Direct-examination confidence</td>
<td>.43</td>
<td>.37</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Cross-examination accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.34</td>
<td>.47</td>
<td>.41</td>
<td>.77</td>
</tr>
<tr>
<td>5. Cross-examination believability</td>
<td>.33</td>
<td>.41</td>
<td>.49</td>
<td>.73</td>
<td>.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Cross-examination confidence</td>
<td></td>
<td>.64</td>
<td>.55</td>
<td>.69</td>
<td>.86</td>
<td>.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Overall believability</td>
<td>.49</td>
<td>.64</td>
<td>.55</td>
<td>.69</td>
<td>.86</td>
<td>.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Overall confidence</td>
<td>.40</td>
<td>.49</td>
<td>.65</td>
<td>.65</td>
<td>.67</td>
<td>.81</td>
<td>.72</td>
<td></td>
</tr>
</tbody>
</table>

Note. All correlations are significant at p < .01.

Jurors were highly discriminating in the way that they evaluated the eyewitnesses. There was considerable variance across the 42 witnesses in the ways that participant jurors perceived the confidence, believability, and accuracy of the witnesses. Furthermore, as evidenced in Experiment 1, there was good interjudge reliability as to which witnesses were thought to be the most confident, believable, and accurate.

There is some cause for concern, perhaps, that participant jurors found the 8-year-old witnesses to be as credible as the 12-year-old and adult witnesses in their cross-examination testimony when, in fact, the 8-year-olds were significantly less accurate. This is yet another example of the difficulty that participant jurors seem to have in evaluating the accuracy of eyewitness testimony (e.g., Wells, 1993; Wells & Leippe, 1981; Wells et al., 1979). In part, the difficulty stems from the fact that eyewitness confidence does not closely track eyewitness accuracy (e.g., Brigham & Cairns, 1988; Leippe, Wells, & Ostrom, 1978; Parker & Caranza, 1989; Wells & Murray, 1983).

Table 3
Perceived Credibility of Participant Witnesses: Study 2

<table>
<thead>
<tr>
<th>Testimony presentation mode</th>
<th>Still frame</th>
<th>Full testimony</th>
<th>t(84)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent measure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct-examination confidence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-year-olds</td>
<td>3.47</td>
<td>1.21</td>
<td>4.40</td>
</tr>
<tr>
<td>12-year-olds</td>
<td>5.20</td>
<td>1.32</td>
<td>4.07</td>
</tr>
<tr>
<td>Adults</td>
<td>4.43</td>
<td>1.14</td>
<td>4.47</td>
</tr>
<tr>
<td>Cross-examination accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-year-olds</td>
<td>3.40</td>
<td>1.30</td>
<td>5.07</td>
</tr>
<tr>
<td>12-year-olds</td>
<td>4.27</td>
<td>1.49</td>
<td>4.40</td>
</tr>
<tr>
<td>Adults</td>
<td>3.53</td>
<td>1.06</td>
<td>3.80</td>
</tr>
<tr>
<td>Cross-examination confidence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-year-olds</td>
<td>3.00</td>
<td>1.13</td>
<td>4.93</td>
</tr>
<tr>
<td>12-year-olds</td>
<td>4.53</td>
<td>1.85</td>
<td>3.73</td>
</tr>
<tr>
<td>Adults</td>
<td>3.73</td>
<td>1.22</td>
<td>3.67</td>
</tr>
<tr>
<td>Cross-examination believability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-year-olds</td>
<td>4.02</td>
<td>1.28</td>
<td>5.70</td>
</tr>
<tr>
<td>12-year-olds</td>
<td>5.20</td>
<td>1.69</td>
<td>4.00</td>
</tr>
<tr>
<td>Adults</td>
<td>4.00</td>
<td>1.16</td>
<td>4.20</td>
</tr>
<tr>
<td>Overall confidence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-year-olds</td>
<td>3.40</td>
<td>1.39</td>
<td>4.93</td>
</tr>
<tr>
<td>12-year-olds</td>
<td>5.20</td>
<td>1.27</td>
<td>4.00</td>
</tr>
<tr>
<td>Adults</td>
<td>4.00</td>
<td>.85</td>
<td>4.20</td>
</tr>
<tr>
<td>Overall believability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8-year-olds</td>
<td>4.00</td>
<td>1.31</td>
<td>5.13</td>
</tr>
<tr>
<td>12-year-olds</td>
<td>4.93</td>
<td>1.39</td>
<td>4.47</td>
</tr>
<tr>
<td>Adults</td>
<td>4.33</td>
<td>1.11</td>
<td>4.60</td>
</tr>
</tbody>
</table>

Note. Ratings were made on a 7-point scale ranging from 1 (not at all) to 7 (extremely). T test compares still frame and full testimony.
* p < .05. ** p < .01.
whereas the believability of eyewitness testimony does closely track eyewitness confidence (e.g., Wells et al., 1979). Recent research has suggested that the difficulty might also derive from the malleable nature of eyewitness confidence (Luus & Wells, 1994).

The current data suggest that the low diagnostic value of eyewitness confidence holds for children as well as adults. Indeed, the perceived confidence of the 8-year-olds' testimony under cross-examination was equal to that of adults, whereas accuracy was lower, and the overall believability of the witnesses corresponded closely to the witnesses' overall confidence ($r = .79$). We believe that if the 8-year-olds' confidence under cross-examination had been lower than that of the adults, the participant jurors would have attributed less accuracy and believability to the 8-year-olds' cross-examination testimony.

There are several possible solutions to the problem of cross-examination producing lower accuracy but not lower perceived credibility for the 8-year-old witnesses versus the adult witnesses. One type of solution is to train or educate jurors to be sensitive to the problems that young children have with cross-examination compared with direct-examination testimony. Another possibility is to follow the lead of some jurisdictions in discouraging or disallowing the use of cross-examination with young children and using only the nonleading styles of questioning. We find neither of these proposed solutions satisfactory. The disallowance of cross-examination tends to interfere with the rights of the accused and violates some basic tenets of the adversarial system. The juror-sensitivity proposal suffers from the general concerns expressed by researchers over trying to estimate the accuracy of witnesses rather than improve the accuracy of witnesses. Although we acknowledge recent experimental success in sensitizing jurors to eyewitness problems (e.g., Cutler, Penrod, & Dexter, 1989) and we agree in principle that expert testimony on these matters should be accepted by the courts (see Kassin, Ellsworth, & Smith, 1989), we favor the development of techniques that can improve the accuracy and the completeness of eyewitness testimony (e.g., see Fisher, Geiselman, & Amador, 1989; Saywitz & Snyder, 1993; Warren-Leubecker, Hulse-Trotter, & Tubbs, 1991). Perhaps these children could benefit from a precourtroom testimony session in which they were taught to recognize discrepancies between their memory and external suggestions, thereby allowing them to resist some of the debilitating effects of cross-examination. Our data clearly demonstrate that there is enough variation in perceived credibility within ages to preclude the assumption that any one person could represent the perceived credibility of that age group.

**Conclusion**

We argue that research on the perceived credibility of child versus adult eyewitnesses must use designs in which eyewitnesses are a sampled variable within age levels and the eyewitnesses are testifying in their own words. Such designs are necessary to capture the natural variance
within ages and thereby avoid the unwarranted assumption that a single child is representative of all children of that child's age.

In general, our results support the contention of Leippe and Romanczyk (1989) that there is an age-based bias or stereotype against the child eyewitness, but when observers view the actual testimony of the children, the bias is suspended. The reader is cautioned, however, to consider several limitations to this work for purposes of generalization.

First, our conclusions are limited to the boundaries of the ages that were used in these studies. Had we used 4-year-old or 6-year-old children, for example, we might have found considerable differences between children of such young age and adult eyewitnesses, even with the provision of their full testimony (see Leippe et al., 1992).

Second, we acknowledge several differences between actual courtrooms and the setting in which our participants were questioned. Recent research has suggested that children may be less accurate in testifying under conditions with high forensic realism rather than a laboratory setting (e.g., Hill & Hill, 1987; Saywitz & Nathan, 1993; Tobey & Goodman, 1992). Testimony of the children and the adults in this research was taken in a familiar, friendly setting (their own schools) and did not include spectators. Under these conditions, the 8-year-old and adult eyewitnesses were generally perceived to be equally confident and credible in testifying. In an actual courtroom, however, the average 8-year-old may be relatively more nervous and appear less confident than the average adult, which, in turn, might confirm rather than disconfirm an observer’s preconception of the child as being less likely to be correct.

These results provide an answer to the question of the fit between the imagined versus actual eyewitness capabilities of children and adults. Our findings suggest that jurors may enter the courtroom with a negative bias against child witnesses. However, qualities of the witness’s testimony seem to play a more important role than does the witness’s age in judgments of eyewitness credibility. That is, people might imagine children to be incapable of delivering accurate, convincing testimony. However, these preconceived negative views of child eyewitnesses may be quickly discarded on actually observing a child testify.

References


Luus, C. A. E., & Wells, G. L. (1994). The malleability of eye-


Received June 28, 1993
Revision received August 1, 1994
Accepted August 2, 1994