

# The Reliability of Children's Eyewitness Testimony

ELIZABETH DUDLEY WESTON



**WRITER'S COMMENT:** *The final assignment for Ryan Honomichl's PSC 141 Cognitive Development class was one of the first American Psychological Association (APA) style papers I wrote. I dreaded this assignment. I had read many empirical research articles in other psychology classes, but I never felt that I completely understood them. I struggled to find articles that were well organized and understandable at an undergraduate level. I was thrilled to find Dr. Gail Goodman's article because it was so well organized and clear. I was inspired; this was the first time that I completely understood an empirical research article. My paper is the product of many hours, numerous revisions, and several disagreements with my T.A. This was an exercise in writing for the person grading your paper, which can be very difficult for any student to swallow. However, it is a very important lesson for us to learn because it carries over to the "real world." Thank you, Ryan, for all of your support and encouragement!*



—Elizabeth Dudley Weston

**Instructor's Comment:** *This paper by Elizabeth was the final assignment for Psychology 141, "Cognitive Development." The course deals with a variety of issues in early cognition from birth through adolescence, such as perception, language, memory, and reasoning. The assignment was to find two empirical articles in prominent psychology journals and compare and contrast their methods, results, and/or theoretical perspectives. Elizabeth chose to look at the issue of children's eyewitness testimony. Specifically, she looked at findings from two experimental studies regarding the accuracy of children's memory recall. Elizabeth revealed several important variables in the study of this issue—besides the more obvious factor of age differences, gender, abuse status,*

*and question type also played some role—and then discussed the legal and psychological relevance of these findings. Overall, her paper shows a thorough grasp of the issues at hand and an ability to discuss them intelligently.*

—Ryan Honomichl, Psychology Department



### **Abstract**

This paper focuses on two studies that examine age differences in the accuracy of children's eyewitness testimony. Goodman et al. (2001) explore the effects of abuse status (abused vs. nonabused), gender, and age on children's eyewitness memory and suggestibility. Cassel and Bjorklund (1995) examine age differences in suggestibility by comparing six and eight-year-old children to adult college students and consider interview type (positive or negative leading) over three interviews (initial, one week, one month). However, Cassel and Bjorklund (1995) do not consider abuse status or gender. Both experiments address issues related to the reliability of children as eyewitnesses. This paper compares and contrasts the two articles, focusing only on the child participants and does not include the adults in Cassel and Bjorklund (1995).

**T**HE RELIABILITY OF CHILDHOOD EYEWITNESS testimony has both legal and psychological relevance. From a legal perspective, any eyewitness testimony can make a strong impression on the jury, thus influencing the jury's decision about guilt or innocence of the defendant. While intentionally making false statements under oath is a crime, inaccurate recall is not. Frequently, legal professionals are concerned about the validity of children's eyewitness testimony because of the belief that children's memories may be more susceptible to suggestion. Research points to age effects in the dependability of children's eyewitness testimony. One study that investigated the effect of age and past abuse experiences on children's eyewitness testimony demonstrated age effects as well as abuse status effects for three- to ten-year-old children (Goodman, Bottoms, Rudy, Davis, & Schwartz-Kenney, 2001). Another study found age differences for free recall and suggestibility, but not for unbiased cued recall, among six- and eight-year-olds (Cassel & Bjorklund, 1995). This kind of analysis is useful because, as cited in lecture, each year more than 100,000 children testify in court. Understanding how children

remember can help legal professionals to appropriately question children to gain accurate information while creating as little stress for the child as possible.

### **Study 1: Goodman, et al. (2001)**

TAKING ABUSE STATUS, AGE, GENDER, ethnicity, and socioeconomic status (SES) into account, Goodman and colleagues (2001) examined the relationship between childhood abuse and children's eyewitness testimony. The experiment consisted of a closely matched sample ( $n = 70$ ) of abused and nonabused children split into two age groups: three to six years ( $N = 30$ ) and seven to ten years ( $N = 40$ ). The children were matched by abuse status, age, gender, socioeconomic level, ethnicity, and delay interval between the initial session and interview. The study consisted of two sessions: an initial play session followed up with an interview about two weeks later. In the play session, each child individually engaged in a social interaction with an unfamiliar male confederate in which they participated in activities such as blowing bubbles, dressing up in costume and posing for a photograph, reading a story, coloring a picture, and thumb wrestling. After this initial session, the children participated in a mock forensic interview where they were asked about various features of the social interaction, including memory of the confederate's appearance. Children were also asked misleading questions about events that did not happen, including events of an abusive nature. They were asked only about the play session, which was a nonabusive social interaction, and not about any actual abuse they may have experienced outside of the experiment. Goodman and colleagues compared the memory and suggestibility of children who had experienced abuse to children who had never been alleged victims or involved in an abuse investigation.

The researchers made three predictions: first, that older children would be more accurate and less suggestible than younger children; second, that abused children would have lower IQ scores and would display more behavioral disturbances than nonabused children; and third, that abused children would be less accurate and more suggestible regarding the nonabuse-relevant facets of the social interaction.

Goodman and colleagues found that in general the older children answered more accurately and completely than younger children with respect to free recall, and older children recalled significantly more correct information than younger children. Boys gave slightly more accurate

responses than girls, but the mean difference was small and thus may be inconclusive. While age had no effect on the boys' performances, younger girls made more incorrect responses than older girls and more errors than younger boys. Younger boys freely recalled more information than the other groups. Older children were more likely to respond "don't know" than younger children, but in this case neither gender nor abuse status produced significant effects. However, the researchers did find an interaction between abuse and gender: abused boys made more "don't know" responses than did nonabused boys.

The researchers did find the expected differences in IQ and CBCL (Childhood Behavioral Checklist) scores: abused children tended to have lower IQ scores and higher CBCL scores. However, while abused children did not express more negative affect, younger children tended to express more negative affect than older children. The researchers found no significant difference in affect by gender.

With respect to free recall, abuse status did not significantly affect accuracy of recalled information. However, researchers found a three-way interaction with age, gender, and abuse status. Young nonabused boys recalled more information than young abused boys (in fact, they recalled more information than did any of the other groups). But the performance of older boys and of girls of all ages was not significantly affected by abuse status. Overall, abuse status did not significantly affect the accuracy of answers to specific questions. Abused and nonabused children did not differ in accuracy or suggestibility in response to questions pertinent to abusive actions. However, nonabused children were more accurate in answering specific questions and made fewer errors in recognizing the unfamiliar adult in a photo identification task. While there were few significant relations between abuse severity and memory measures, severity in terms of abuse type was significantly related to performance on specific, but not misleading, abuse-related questions. Children who had suffered more severe sexual abuse made more omission errors specific to abuse-relevant questions; children who had suffered more invasive abuse tended to commit more omission errors, i.e., failing to report an abusive action that did occur. Children who were older at the onset of abuse provided more correct information during free recall than those who experienced abuse at a younger age.

When presented with misleading questions, children in the study tended to resist most suggestions about fictitious events. Older children

gave more correct responses than younger children to misleading questions, and boys gave more correct answers than girls. Overall, the children made remarkably few errors in response to the abuse-related queries, in particular to misleading ones.

The children were also asked to identify the male confederate from a target-present photo lineup, which yielded three possible responses: correctly identifying the confederate, incorrectly identifying a different photo as the confederate, and incorrectly stating that the confederate's photograph was not present. Abused children made more than twice as many incorrect identifications as nonabused children. There were no other significant main effects or interactions involving age, gender, or abuse status on the proportion of correct, incorrect, or failed identifications.

While age differences were not found in the photo identification task, the research pointed to age differences in memory abilities and suggestibility. Even the youngest children in the study recalled important information and resisted misleading questions. However, with respect to free recall, specific and misleading questions, older children provided more correct answers than younger children. These age patterns emerged for abused and nonabused children. In general, Goodman et al. (2001) showed that older children recalled more and gave more accurate responses than younger children.

## **Study 2: Cassel and Bjorklund (1995)**

USING METHODS INTENDED TO REPLICATE the general experience that a fact witness could expect in a real case, Cassel and Bjorklund (1995) assessed the effects of being asked leading questions over repeated assessment. They hypothesized that their experiment would support previous research that indicated that, in general, young children's free recall of events tends to be low, but is in most cases accurate. Citing a different study that showed levels of free recall for three- and six-year-old children were fairly high and both groups tended to answer misleading questions correctly, the researchers asserted that in certain circumstances younger children tend to be more suggestible than older children (Ornstein et al. in Cassel & Bjorklund, 1995).

The mock trial technique included methods of questioning and delays that a real witness might experience. The experiment consisted of six-year-old ( $N = 45$ ) and eight-year-old ( $N = 45$ ) children as well

as 70 college-aged adults. (Please recall that this paper looks only at the results for the children and not the adults in this study.) All of the children attended public schools and came from predominantly middle class families. Participants were randomly assigned to one of three treatment levels: control, positive leading, and negative leading. Participants viewed a brief film in which a 14-year-old boy and a nine-year-old girl argued about a bicycle. After the video, the children were given a puzzle to play with for 15 minutes and then free recall was assessed for all three groups. The children were asked to tell the interviewer everything they could remember about the film. If participants hesitated more than ten seconds, the interviewer asked, “Can you remember anything else?” When the participant could not recall more, the interviewer then asked unbiased cued-recall questions. Over the next month, participants were interviewed one or two more times. After one week, only the positive and negative leading groups were interviewed again. After one month, all three groups were asked for free recall again and were asked the same positive and negative leading questions asked at the one week interview in a mock trial setting.

For the initial interview, overall levels of correct free recall increased with age and inaccurate free recall was almost nonexistent for all participants. Total correct recall (free recall and unbiased cued recall) showed significant age differences, with the eight-year-olds recalling more than the six-year-olds. The significant effect of age found at the initial interview remained at both the one week and one month interviews. While some forgetting did occur, analysis revealed no significant age effect. The researchers found no significant interactions.

At the one week interview, both groups of children in the positive condition provided more correct answers than their peers in the negative condition. At the one month interview, the six-year-olds were more likely to correctly answer positive leading questions than the eight-year-olds, and the eight-year-olds were more likely to correctly answer the negative leading questions. In general, patterns of incorrect recall were opposite those of correct recall. Subjects in the negative leading condition had higher incorrect recall than those in the positive leading condition. Six-year-olds rarely rejected positive leading questions, whereas eight-year-olds were about twice as likely as the younger children to reject positive leading questions. In the negative leading condition, the researchers report that younger children were more likely to incorrectly

accept the negative leading questions. These percentages are fairly close (62 percent vs. 57 percent), however, and taking a reasonable margin of error into account, the significance of this difference is unclear. At the one month interview, the older children responded incorrectly to positive leading questions more often than the younger children. Conversely, the younger children responded incorrectly more often than the older children to negative leading questions. At the one week interview, six- and eight-year-olds were about equally likely to respond "don't know" to leading questions. At the one month interview, five percent of the older children responded "don't know" to the leading questions while none of the younger children did. Again, this data is inconclusive.

The researchers separated the questions into four categories: (1) central items (e.g., who owned the bike, did the boy have permission to take the bike); (2) appearance items (e.g., of the people in the video); (3) bicycle items (e.g., bike color); and (4) miscellaneous items (e.g., setting and weather). They analyzed how the children changed their answers from the initial to the one month interview. Free recall of the central items was highest for all groups and was virtually never incorrect. Correct free recall of miscellaneous items was less than for central items but more than appearance and bicycle items, which were at or near floor levels for both groups of children. Both age groups were more likely to change response for a peripheral issue than for a central issue, and in general younger children were more likely than older children to change. The researchers were surprised when bicycle category items were recalled incorrectly more often than any other item in the study and note that this is the type of error that can destroy a witness's credibility if presented in testimony at trial.

## **Research Methods**

WHILE BOTH OF THESE STUDIES addressed the reliability of children's eyewitness testimony using a mock trial setting, their experimental methods differed. Both of these studies attempted to replicate a real eyewitness experience. All participants in Goodman et al. (2001) had the same two-week delay from the first session to the interview session, while participants in Cassel and Bjorklund (1995) had differing time delays—all participants were interviewed after a 15-minute delay, the two experimental groups were interviewed again after one week, and all three groups participated in the mock trial after one month. These two

studies addressed different situations in which a child may be an eyewitness: one that directly involved the child, and one that the child simply observed. The children in Goodman et al. were active participants in the interaction, whereas in Cassel and Bjorklund the children observed a recorded interaction between two people. Research has shown that children who experience an event firsthand report more accurately on the event than children who only observe an event. When children directly experience something, they tend to think more deeply about it and thus recall more (Siegler, 2005, p. 231). Abused children may be more likely than nonabused children to have to testify about something that happened to them. On the other hand, abused and nonabused children are about equally likely to have to testify about an event that they witnessed by chance but that does not immediately involve them. Consequently, abuse status is relevant in Goodman et al., but it may not be in Cassel and Bjorklund.

### **Sample**

GOODMAN ET AL. (2001) TOOK VARIOUS FACTORS into account, such as age, gender, abuse status, IQ, and ethnicity, whereas Cassel and Bjorklund (1995) looked only at age. Goodman and colleagues focused entirely on two different age ranges of children, while Cassel and Bjorklund compared two specific ages of children to college-aged adults. The age groups in both studies are somewhat comparable, but Goodman et al. looked at a broader range of ages. The mean age of children in Goodman et al. was seven years and six months, which is closer to the older group of children in Cassel and Bjorklund. However, it is very likely that this is at least in part because Goodman and colleagues had more children in the older (N = 40) than younger (N = 30) category. This may raise concerns when comparing the “younger” children and “older” children from each experiment. Knowing the mean age for each group in Goodman et al. might lead to clearer direct comparisons between the two studies.

That Goodman et al. looked at a more varied range of ages may make their study more generalizable. It is unclear why Cassel and Bjorklund chose to limit their sample to children in kindergarten (six-year-olds) and second grade (eight-year-olds). Granted, a six-year-old and an eight-year-old are at different stages of cognitive development, but it might have been appropriate to consider a broader age range or even a third group of children. However, perhaps financial or other constraints limited the size and scope of their sample.

From a Piagetian perspective, the two studies compared preoperational and concrete operational children. According to Piaget, preoperational children (age two to seven years) understand the world through language and mental imagery, whereas concrete operational children (age seven to twelve) are beginning to understand the world through logical thought and categories. Notice that the mean age of Goodman et al. (2001) fell squarely between the preoperational and concrete operational stages, while in Cassel and Bjorklund (1995) the “younger” children were more towards the older end of the preoperational stage and the “older” children were closer to the beginning of the concrete operational stage.

A three-year-old is quite different cognitively from a six-year-old, and a seven-year-old from a ten-year-old, so it seems reasonable to assume that even within these two groups a great deal of variation in cognitive development existed. On the other hand, the six- and eight-year-olds in Cassel and Bjorklund fall neatly into Piaget's stages, preoperational and concrete operational, respectively. It is reasonable to assert that the two groups in Cassel and Bjorklund would be more homogeneous with respect to cognitive capacity than those in Goodman et al.; for example, there should be less cognitive variation in a group of all six-year-olds than in a group of children ranging from ages three to six years. The fact that Goodman et al. included children as young as three and as old as ten indicates that the range of cognitive capacities is far greater than in Cassel and Bjorklund.

Goodman and colleagues' (2001) sample was ethnically and socioeconomically diverse, whereas the Cassel and Bjorklund (1995) sample was predominantly middle class and they did not specify the ethnicities of the participants. However, since Goodman et al. found no significant effects of ethnicity on any of the dependent measures, perhaps it is reasonable to assume that the ethnic composition of the Cassel and Bjorklund sample may be irrelevant. Goodman and colleagues found a significant, but weak, negative correlation ( $r = -0.24$ ,  $p < 0.05$ ) between socioeconomic status (SES) and free recall. Thus, perhaps SES is by and large irrelevant in Cassel and Bjorklund.

## **Recall**

**BOTH EXPERIMENTS REVEALED SIGNIFICANT age effects; overall, older children recalled significantly more than younger children. Additionally, older children recalled more accurate information than younger children,**

but in general most of what the children recalled was accurate. Cassel and Bjorklund (1995) found that incorrect free recall was almost nonexistent for all participants, while Goodman et al. (2001) found that older children gave more correct free recall responses than younger children. Several factors may contribute to the differing age-related memory abilities. Generally speaking, older children are more cognitively advanced than younger children in many areas, including source monitoring, language abilities, and information processing. Furthermore, older children tend to have a larger knowledge base to draw upon than younger children (Brainerd & Reyna; Case; Chi; Fischer; Fivush; Johnson, et, al.; Lindsay; Nelson, in Goodman et al., 2001).

From a Piagetian perspective, preoperational children still see the world from a largely egocentric perspective, and they also tend to focus on a single aspect of stimuli. From this perspective, it makes sense that in general younger children would recall less than older children, because they may not actually take in as much as an older child, who can focus on multiple aspects of stimuli.

### **Suggestibility**

BOTH STUDIES FOUND AGE EFFECTS for suggestibility. Cassel and Bjorklund (1995) found differences in suggestibility for both age groups depending on the type of information being recalled. Items considered to be central to the case—for example, whether or not the teenage boy had permission to take the bicycle from the female child—were more likely to be recalled correctly by all of the participants. According to the researchers, 79% of the six-year-olds and 91% of the eight-year-olds initially gave the “correct” answer to the permission question. Subsequently, while only seven percent of the eight-year-olds changed their original answer to the permission question in response to a leading question, 42 percent of the six-year-olds changed their original answers following a leading question. The important issue was that the younger children tended to accept leading questions significantly more than did the older children.

However, this permission question is of concern because it seems subjective rather than having a clear right or wrong answer. Whether or not the older boy had permission to take the bike from the younger girl is a judgment call. Perhaps this lends realism to the experiment; expecting a “witness” to make such a judgment call is reasonable because this may be expected in a real court setting. But how could the viewer possibly

know the “right” answer? The bike in question was reported to be a common BMX-type bicycle with red tires, but the researchers did not specify whether it was actually a boy's or girl's bicycle, which is often very obvious in children's bikes. Furthermore, while children are less egocentric by the age of six, a six-year-old may not be entirely capable of viewing an event like this neutrally. A child's previous experiences in similar situations may color his or her judgment about the permission question. Perhaps this is trivial and the real issue is not whether the answer was originally “correct,” but rather whether or not the participant subsequently changed his or her answer.

In light of controversy surrounding the credibility of children's eyewitness testimony, one very important finding in Goodman et al. (2001) was that abused children were no more suggestible than nonabused children. Goodman and colleagues also found that even the youngest children in the study resisted misleading questions, although older children were more likely to answer misleading questions correctly than younger children. Child abuse is an all too common problem in our society, and many investigations rely on the abused child's testimony. Clearly, the credibility of child eyewitnesses is critical to legal and criminal issues surrounding such cases.

### **Psychological and Legal Relevance**

BOTH OF THESE STUDIES ADDRESSED important psychological and legal issues involving children's eyewitness testimony, and both reasonably recreated a realistic mock trial setting. Goodman et al. (2001) considered individual differences in the children by taking numerous factors into account, whereas Cassel and Bjorklund (1995) considered only age. Additionally, other factors such as gender and SES that Goodman et al. considered may be immaterial in Cassel and Bjorklund. For example, gender is relevant in Goodman et al. because girls are about twice as likely to be abused as boys (Finkelhor in Goodman et al., 2001), but boys and girls may be about equally likely to witness one child taking something from another child. Thus, gender may not be relevant in Cassel and Bjorklund.

Goodman and colleagues suggested that future researchers investigate additional measures, such as family stability, post traumatic stress disorder, attachment, and dissociation. Goodman et al. also recommended investigating the effects of repeated questioning and false memory effects

in maltreated children, but acknowledged that ethical concerns should first be carefully evaluated. Previous research has shown that repeated questioning can lead children to answer the same question differently. Sometimes the child recalls more when asked the same question more than once. However, it is possible that young children sometimes modify their answers hoping to please the interviewer (Siegler, 2005, p. 231).

Being an eyewitness can be a very stressful experience for anyone, regardless of age. While realism in an experiment can increase the ecological validity of a study, every effort must be made to protect the participants from long-term adverse effects. Both of these studies point to age-related differences in children's eyewitness testimony. Cassel and Bjorklund note that younger children encode information differently than older children. Younger children tend to rely on verbatim representations, while older children rely on gist representations (Siegler, 2005, p. 228). Young children's tendency to rely on verbatim traces may in part explain their greater susceptibility to suggestion than older children (Ceci & Bruck, in Cassel & Bjorklund, 1995). Furthermore, taking factors such as the child's general level of cognitive development into account, future researchers can help to develop age-appropriate methods of questioning that can obtain accurate information while minimizing trauma for the child.



## References

- Cassel, W. S., & Bjorklund, D. F. (1995). Developmental patterns of eyewitness memory and suggestibility: An ecologically based short-term longitudinal study. *Law and Human Behavior, 19* (5), 507–32.
- Goodman, G. S., Bottoms, B. L., Rudy, L., Davis, S. L., & Schwartz-Kenney, B. M. (2001). Effects of past abuse experiences on children's eyewitness memory. *Law and Human Behavior, 25* (3), 269–98.
- Siegler, R.S., & Alibali, M.W. (2005). *Children's thinking* (4<sup>th</sup> ed.). New Jersey: Prentice Hall.

