EFFECTS OF INTERVIEWER BEHAVIOR ON ACCURACY OF CHILDREN'S RESPONSES

JESSICA SPARLING, DAVID A. WILDER, JENNIFER KONDASH,
MEGAN BOYLE, AND MEGAN COMPTON

FLORIDA INSTITUTE OF TECHNOLOGY

Previous research has shown that certain interviewer behaviors can evoke inaccurate answers by children. In the current study, we examined the effects of approving and disapproving statements on the accuracy of 3 children's answers to questions in an interview (Experiment 1). We then evaluated 3 questioning techniques that may be used by interviewers during a forensic interview in which a child provides eyewitness testimony (Experiment 2). All participants responded with more inaccurate answers when approving statements followed inaccurate information and disapproving statements followed accurate information in Experiment 1. During Experiment 2, 1 participant responded most inaccurately when she was requestioned after providing an initial answer, whereas the remaining 2 participants responded most inaccurately when the interviewer provided cowitness information and suggestive questions.

Key words: accuracy of children's responses, eyewitness testimony, forensic interviews, interviewer behavior

_____

Evaluation of the accuracy of children's reports has become an important topic (Saywitz & Camparo, 2009). Some research on child interviews has focused on how the behavior of the interviewer influences children's accuracy. Garven, Wood, Malpass, and Shaw (1998) suggested that improper interviews have the potential to evoke false information from children and described four broad categories of improper interview techniques: suggestive questioning (e.g., asking "Did he touch her on the bottom?" when a child has acknowledged touching but has not mentioned any inappropriate touching), influence (e.g., telling a child what a cowitness has said or has seen about a topic), reinforcement and punishment (i.e., delivery of approving and disapproving statements contingent on a child's report), and removal from direct experience (e.g., inviting a child to speculate about what might have occurred).

Garven et al. (1998) examined the effects of suggestive questions, cowitness information, approving and disapproving statements, and requestioning (i.e., re-presenting a question after an initial answer) on children's accuracy in interviews. One week after a classroom visit by a graduate student, children were exposed to all five of the above techniques or only to suggestive questions during interviews about the visit. Children exposed to the five interview techniques were more likely to answer questions inaccurately; however, it was not possible to determine which of the five techniques had the greatest effect on the children's responding. In a follow-up study, Garven, Wood, and Malpass (2000) examined the effects of approving statements, cowitness information, and suggestive questioning on children's accuracy. The children responded inaccurately most often when the interviewer delivered approving statements contingent on the child providing inaccurate information.

Despite the importance of this topic, few studies have examined these interviewing techniques individually, and only one study (which focused on suggestive questioning) has been conducted at the individual level using single-

_____

subject methodology (Doepke, Henderson, & Critchfield, 2003). Repeated measurement of the dependent variable, which is characteristic of single-subject methodology, may be particularly relevant for this topic; children who participate as eyewitnesses in legal investigations may be asked to recount events multiple times (Zaragoza, Graham, Hall, Hirschman, & Ben-Porath, 1995). Thus, the purpose of the current study was to examine further the effect of interviewing techniques on children's accuracy. In the first experiment, we evaluated the effects of approving and disapproving statements. The effects of suggestive questions, cowitness information, and requestioning were evaluated in a second experiment.

METHOD

Participants, Setting, and Materials

Three children participated: Matthew (4 years old), Owen (5 years old), and Abby (8 years old) . All children had age-appropriate language skills, had no disabilities or psychiatric diagnoses, and were in general education classrooms at their schools. All sessions took place at a university in a room that contained a computer (to watch videos) and a table and two chairs. Videos were obtained from various online resources and consisted of clips of children's cartoons (e.g., Scooby Doo) that were 3 to 5 min in length. Up to 20 questions and corresponding correct answers based on events in each video were prepared in advance by a graduate student. Videos were randomly assigned to the various conditions of the studies; that is, each video had an equal chance of being used in each condition of each study. Tto ensure that repeated exposure to a video did not affect responding, participants watched each video only once.

Response Measurement

The dependent variable was the accuracy of participants' responses to interviewer questions. Participants were interviewed in sessions that took place over several days. Accurate responses were defined as a "yes" or "no" response to a question that required the corresponding answer. Inaccurate responses were defined as a "yes" or "no" response to a question for which the opposite answer would be appropriate. If a participant answered "I don't know" to a question, that question and answer were excluded from the analysis (one to two questions per session were excluded for each participant). In Experiment 1, reinforcement was defined as saying "good" and smiling. Punishment was defined as saying "that's not right" and frowning. A neutral response was defined as a quick nod and saying "okay" with no vocal inflection. General questions, which were used in both experiments, were defined as yes or no questions that contained information about events in the videos (e.g., asking if Scooby fell, if he had fallen). In Experiment 2, suggestive questions were defined as yes or no questions that contained information about events that did not occur in the videos (e.g., asking "Didn't Scooby fall?" when he had not). Cowitness information was defined as a question preceded by the statement "Someone told me that …" Requestioning was defined as an interviewer repeating a question twice after an answer had been given. In this condition, the participant's answer to the final question was scored.

Trained graduate students scored responses (either accurate or inaccurate) on data sheets. Interobserver agreement data were collected for at least 33% of responses for each condition per participant by having a second observer collect data from video or audio recordings of interviews. Exact agreement was calculated on a trial-by-trial basis by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100%. Agreement ranged from 95% to 100% in Experiment 1 and 94% to 98% in Experiment 2. Independent variable integrity was determined by recording the number of correct and

incorrect interviewer responses or deliveries (i.e., delivery of reinforcement, punishment, or a neutral response in Experiment 1 and delivery of the question in the general, suggestive, cowitness, or requestioning format in Experiment 2), dividing correct responses by total responses, and multiplying by 100%. Independent variable integrity ranged from 94% to 98% during Experiment 1 and 97% to 100% for Experiment 2.

Procedure

Experiment 1. At the start of Experiment 1, participants were told to watch all videos closely because they would be asked some questions. For each session, participants were shown the video (described above) and were subsequently interviewed by the experimenter. Interviews were preceded by the direction "I am going to ask you some questions. Please answer 'yes' or 'no,' and do the best you can." Four different conditions were implemented: control, reinforcement for inaccurate responding, punishment for accurate responding, and reinforcement plus punishment. The interviewer, who was the same person throughout both experiments, wore a different colored shirt for each condition to assist participants to discriminate between conditions. The effect of the different consequences on response accuracy was compared using a multielement design.

General questions were asked during all conditions, with a total of 12 questions asked per session for Matthew and Abby (due to time constraints), and 20 questions per session for Owen. In the control condition, reinforcement was delivered for each response, independent of the accuracy of the response. Thus, one would expect to see lower (but not nonexistent) levels of inaccurate responses in this condition than in the other conditions. In the reinforcement for inaccurate responding condition, all accurate answers were followed by a neutral response and inaccurate answers were followed by reinforcement, as defined above. In the punishment for accurate responding condition, all accurate

answers were followed by punishment, as defined above; inaccurate answers were followed by a neutral response, as defined above. In the reinforcement plus punishment conditions, reinforcement was delivered contingent on inaccurate answers and punishment was applied to accurate answers.

Experiment 2. Procedures for Experiment 2 were identical to those of Experiment 1, with the exception of the techniques evaluated. That is, participants were shown the video and were subsequently interviewed by the experimenter. Interviews were preceded by the direction "I am going to ask you some questions. Please answer 'yes' or 'no,' and do the best you can." Three different interviewer techniques were evaluated in Experiment 2: (a) suggestive questions, (b) cowitness information, and (c) requestioning. During baseline, the interviewer asked general questions and responded neutrally to all of the child's answers. During the suggestive questions condition and cowitness information conditions, the interviewer asked eight suggestive questions per session for Matthew and Abby (due to time constraints) and 14 suggestive questions per session for Owen. In the baseline and requestioning conditions, the interviewer asked 12 questions per session for Abby and Matthew and 20 for Owen. Effects of interview techniques were evaluated using an ABACA-DABACAD reversal design for Owen and Matthew and an ABCDABCD design for Abby (due to time constraints), in which A was baseline, B was the suggestive questions condition, C was the cowitness information condition, and D was the requestioning condition.

RESULTS AND DISCUSSION

Results of Experiment 1 are depicted in Figure 1 (left). The highest levels of inaccuracy occurred during the reinforcement plus punishment conditions for all participants. In addition, the control condition produced relatively low levels of inaccuracy for all participants. Matthew responded inaccurately to a mean of

33% (range, 8% to 50%) of questions in the reinforcement plus punishment condition and a mean of 8% (range, 0% to 17%) of questions in the control condition. Owen responded inaccurately to a mean of 53% (range, 37% to 65%) of questions in the reinforcement plus punishment conditions, and a mean of 20% (range, 10% to 39%) of questions in the control condition. Abby responded inaccurately to a mean of 58% (range, 25% to 92%) of questions in the reinforcement plus punishment condition and a mean of 17% (range, 8% to 25%) of questions in the control condition.

Results of Experiment 2 are depicted in Figure 1 (right). Matthew responded most inaccurately in the suggestive questioning condition ($M = 81\%$; range, 50% to 100%) and cowitness information condition ($M = 69\%$; range, 38% to 100%). Matthew's percentages of inaccurate answers in the requestioning and baseline conditions were 4% (range, 0% to 10%) and 3% (range, 0% to 10%), respectively. Owen was most inaccurate in the cowitness information condition ($M = 52\%$; range, 14% to 71%) and the suggestive questioning condition ($M = 49\%$; range, 21% to 100%). Owen's percentages of inaccurate answers in the requestioning and baseline conditions were 6% (range, 0% to 11%) and 15% (range, 0% to 61%), respectively. Abby's highest levels of inaccurate responding occurred in the requestioning condition ($M = 87\%$; range, 57% to 100%), relative to suggestive questioning ($M = 20\%$; range, 0% to 50%), co-witness information ($M = 25\%$; range, 12% to 63%), and baseline conditions ($M = 53\%$; range, 17% to 100%).

These results suggest that both antecedents (i.e., the structure of questions) and consequences (i.e., reinforcement and punishment) may influence the accuracy of children's responding to an interviewer's questions. Further, the differential responding observed across participants suggests that some children may be more or less sensitive to specific interviewer techniques. This is important, because the majority of prior studies on this topic have not been conducted in a manner that allows such analysis. Finally, in previous studies neither suggestive questioning (Garven et al., 1998) nor providing cowitness information (Garven et al., 2000) were found to increase inaccuracy substantially. In contrast, results of the current study suggest that both of these techniques may increase inaccuracy for some children. The methodological differences between the current study and previous research may account for this finding.

Of course, the methodology employed in the current study is different from a forensic interview in which a child provides an eyewitness testimony. In the current study, children were asked about events they observed in cartoons. In actual forensic interviews, children are often asked questions about potentially traumatic events they observed in real life. Nevertheless, for ethical and practical reasons, questions about fictional characters and events are standard in research on eyewitness testimony (Ceci & Bruck, 1995). Also, although the dichotomous questions and the interview techniques employed in this study are not typical of forensic interviews, they are sometimes used (Garven et al., 1998; Saywitz & Camparo, 2009). An additional limitation is the small sample size; future research should employ more participants.

Children in this study ranged in age from 4 to 8 years. It is likely that some developmental differences in language skills exist among this age range that might affect accuracy. Indeed, forensic interview guidelines reflect this; recommendations for younger children call for more specific questions than for older children (Saywitz & Camparo, 2009). In Experiment 2 of the current study, the oldest participant responded most inaccurately when the interviewer re-questioned her after an initial answer, whereas the younger children responded most inaccurately in other conditions. Future re-
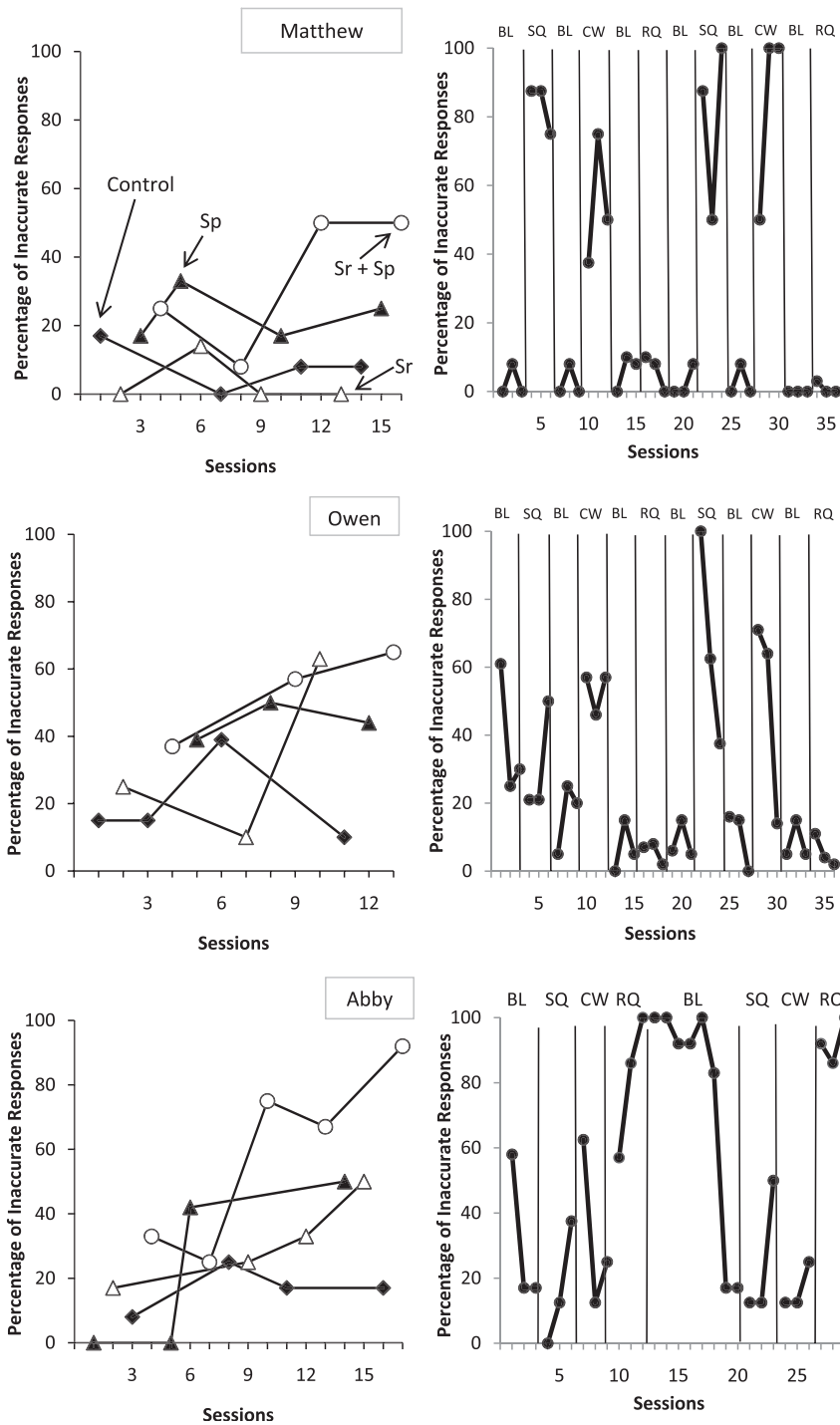
Figure 1. Percentage of inaccurate responses during control, reinforcement for inaccurate responding (Sr), punishment for accurate responding (Sp), and reinforcement for inaccurate plus punishment for accurate responding (Sr + Sp) conditions for Matthew, Owen, and Abby in Experiment 1(left). Percentage of inaccurate responses during baseline (BL), suggestive questions (SQ), cowitness information (CW), and requestioning (RQ) conditions for Matthew, Owen, and Abby in Experiment 2 (right).

search should investigate additional age-related differences in sensitivity to interviewer techniques.

REFERENCES

Ceci, S., & Bruck, B. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. Washington, DC: American Psychological Association.

Doepke, K. J., Henderson, A. L., & Critchfield, T. S. (2003). Social antecedents of children's eyewitness testimony: A single-subject experimental analysis. *Journal of Applied Behavior Analysis*, *36*, 459–463.

Garven, S., Wood, J. M., & Malpass, R. S. (2000). Allegations of wrongdoing: The effects of reinforcement on children's mundane and fantastic claims. *Journal of Applied Psychology*, *83*, 38–49.

Garven, S., Wood, J. M., Malpass, R. S., & Shaw, J. S. (1998). More than suggestion: The effects of interviewing techniques from the McMartin preschool case. *Journal of Applied Psychology*, *83*, 347–359.

Saywitz, K. J., & Camparo, L. B. (2009). Contemporary child forensic interviewing: Evolving consensus and innovation over 25 years. In B. Bottoms, C. Najdowski, & G. Goodman (Eds.), *Children as victims, witnesses, and offenders: Psychological science and the law* (pp. 102–127). New York: Guilford.

Zaragoza, M. S., Graham, J. R., Hall, G. C. N., Hirschman, R., & Ben-Porath, Y. S. (1995). *Memory and testimony in the child witness*. Thousand Oaks, CA: Sage.