Improving Children's Person Identification

Beth M. Schwartz-Kenney
Randolph-Macon Woman's College

Bette L. Bottoms
University of Illinois at Chicago

Gail S. Goodman
University of California, Davis

The accuracy of children's person identification on a recall task was compared to their accuracy on specially developed feature lineup tasks. In two experiments, children participated in a play session with a stranger and subsequently were interviewed about the stranger's appearance. Each child was first asked to recall appearance features (e.g., height, age, skin color) and then was asked to identify the features using target-present lineups. In Experiment 1, 48 3- to 6-year-olds used feature lineups immediately after the play session. In Experiment 2, 48 4- to 9-year-olds used modified feature lineups that included a "don't know" option after a 2-week delay. Children's identification accuracy was significantly related to age and type of identification task. In Experiment 1, use of the feature lineups increased children's correct responses but also increased their incorrect responses. Use of the modified feature lineups in Experiment 2 led to increased accuracy with mixed results concerning commission errors. In both experiments, older children were more accurate than younger children.

In 1986, Jones and Krugman detailed the case of a 3-year-old child who was kidnapped from her front yard, sexually assaulted, and left to die in a remote mountain area. By chance, she was found and rescued, but what made the case particularly remarkable was her ability to positively identify her abductor from a lineup. His confession corroborated her identification.

Children often must give reports about past events involving strangers. They are increasingly likely to witness criminal events such as shootings and robbery, and many children are themselves physically and sexually abused each year (e.g., Finkelhor, Hotaling, Lewis, & Smith, 1990). Although stranger-perpetrated child abuse is less common than nonstranger abuse, it does occur with sufficient frequency to warrant concern over children's identification abilities. In contrast to the successful investigation and prosecution detailed by Jones and Krugman (1986), investigations and adjudication of cases can be quite difficult when key evidence is children's reports about a stranger's identity.

Little is known about young children's ability to report pertinent information about the appearance of formerly unknown people, and even less is known about how to obtain accurate descriptive reports from children. Although some researchers have begun to address ways to improve children's eyewitness memory (e.g., Saywitz & Snyder, 1993), few have considered the specific problem of improving children's person identification. In the present research, we developed and investigated the effectiveness of a technique designed to obtain descriptive information about strangers from children. We report here the theory behind the development of our technique, two studies investigating its effectiveness, and implications for future research in this important area of child witness research.

TECHNIQUES TYPICALLY USED TO AID PERSON IDENTIFICATION

Although numerous types of information are crucial in a forensic investigation, Chance and Goldstein (1984) note that the identification of a perpetrator is "often the definitive evidence upon which arrests are made and convictions obtained" (p. 69). Several techniques currently are used in our criminal justice system to obtain person identification information from

CHILD MALTREATMENT, Vol. 1, No. 2, May 1996 121-133

adult witnesses. We consider each in turn, discussing its potential for use with children.

Free Recall

The simplest technique, and one that may be the least contaminating or suggestive, is asking witnesses largely open-ended recall questions such as "What did the person look like?" (Dent, 1991) without the presence of cues. Although adults may provide useful descriptive information in response to such a query, children's unprompted narratives are likely to be quite limited (Davies, Tarrant, & Flin, 1989; Goodman, Rudy, Bottoms, & Aman, 1990). Importantly, however, the few details that children do provide are usually, although not always, accurate. For example, in a study by Goodman and colleagues (1990), 4- to 7-year-old children who interacted with a stranger later recalled little (not much more than one piece of information on average) in response to the question, "What did the man look like?" Brigham, Van Verst, and Bothwell (1986) found that even 10- to 17-year-olds have difficulty reporting age and that 10-year-olds have trouble reporting weight. This is not surprising given the findings of Flin and Shepherd (1986) that even adults are not very accurate when asked questions about the height and weight of a briefly encountered stranger.

Children's ability to provide only small amounts of information in response to a recall task is due in part to their limited ability to use retrieval strategies (Goodman, 1984). Numerous studies have demonstrated that understanding and use of memory strategies, such as the unprompted use of self-generated retrieval cues, increase with age (Brown, 1979; Kobasigawa, 1974). In addition, free recall tasks demand fairly advanced language production abilities that are not well developed in young children (Kail, 1990; Whitehurst, 1982). Therefore, a free recall task, which by definition does not provide retrieval cues and relies on a verbal response, is likely to generate an incomplete account of a witnessed event from young children and may underestimate children's actual memory of an event. As suggested by Mandler (1990), in some circumstances "younger children recall as much as older children, but it takes a lot more work to extract the information from them" (p. 325).

Cued Recall: "Prop" Kits

To increase the amount of pertinent identification information elicited from adults, authorities often supplement recall questions with specific cueing techniques such as the use of prop or cue kits like the "Photofit" and "Identi-Kit." These kits contain representations of various facial feature "cues" (e.g., noses, mouths, eyes) that witnesses are asked to use in constructing a composite of the target person's face. From a cognitive perspective, such an approach is theoretically justifiable because it changes the task of remembering details of a face from recall to recognition. It is a robust finding that recognition memory generally is superior to recall (e.g., Bahrick, Bahrick, & Wittlinger, 1975; Meyer & Hilterbrand, 1984). Further, the presence of cues increases similarity between encoding and retrieval context, thus enhancing retrieval of memories (Tulving, 1983).

The use of props or objects representative of those present at a target event can, under some conditions, improve children's reports for past events (Goodman & Aman, 1990; Piaget & Inhelder, 1973; Pipe, Gee, & Wilson, 1993; Price & Goodman, 1990; Saywitz, Goodman, Nicholas, & Moan, 1991). Children, like adults, are better at recognition tasks than they are at recall tasks. Children may also benefit from the reduced language demands of a recognition task as compared to those of a purely recall task. Recognition tasks that expose adults or children to information that is not part of an original experience, however, may also increase the possibility of suggestibility or misinformation effects (e.g., Loftus, 1979). That is, introducing false information may lead witnesses to incorporate incorrect information into their reports of past events or, in this case, reports of a stranger's appearance. This is of particular concern with young children (Saywitz et al., 1991). Thus, although researchers have demonstrated that children's accuracy on memory tasks can be improved when the task relies on recognition rather than recall of to-be-remembered information, caution is warranted when the cues represent information that was not part of the original event.

Are cue kits actually helpful for adult or child witnesses? The effectiveness of the Photofit, a kit that includes actual photographs of facial features, was examined by Flin, Markham, and Davies (1989). Adults' identification performance was better than that of 8- and 11-year-olds, but the overall accuracy for all age groups was low, a finding consistent with other research assessing subjects' ability to use composite construction tasks (Davies, 1981; Laughery & Fowler, 1980). In fact, Flin and colleagues found that subjects of all ages were significantly more accurate at facial descriptions when simply asked to verbally describe the person than they were when using the Identi-Kit.

Comish (1987) examined the issue of suggestibility by determining whether use of the Identi-Kit influenced subsequent recognition performance. Interestingly, she found that after subjects had built a composite using the Identi-Kit, they made a greater

number of false alarm errors during a subsequent recognition task. The subjects' task was to choose an original composite face presented earlier from a lineup series of five composites. Comish found that subjects chose distractors from the lineup that contained characteristics similar to the composite they built with the Identi-Kit rather than to the original composite face. She concluded that use of the Identi-Kit increased errors when performing the lineup identification task, and these errors were consistent with previous errors made when using the Identi-Kit.

Thus, in contrast to the conventional wisdom of some police departments, the task of face reconstruction using complex cue kits may not always lead to more accurate person recall for either adults or children and in fact could lead witnesses to incorporate incorrect information into their person identification reports.

Lineups

In authorities have enough evidence to suspect a particular person of committing a crime, they may use a direct test of a witness's recognition memory—live or photographic lineups. Rather than asking a witness to provide descriptive information about a person or to choose from large numbers of facial features, a witness is asked to view a lineup composed of live subjects or photographs that do (target present) or do not (target absent) include the target suspect. Suspects may be presented sequentially or, more commonly, simultaneously.[1] A large body of psychological research has considered adults' ability to respond accurately to lineups (see Wells, 1993, and Wells & Seelau, 1995, for reviews). More recently, researchers have begun to examine children's accuracy in responding to lineups. A consistent finding is that children are particularly prone to make false identifications when using target-absent lineups (King & Yuille, 1986; Parker & Carranza, 1989). For example, in a study by King and Yuille, 8- to 11-year-olds made false identifications 74% of the time on a target-absent lineup when attempting to identify a stranger. Children's commission errors in response to target-absent lineups may be caused by demand characteristics inherent in the task (Goodman, Bottoms, Schwartz-Kenney, & Rudy, 1991; King & Yuille, 1986). Specifically, children may incorrectly assume that the best response to a lineup task is to make a choice, even if the choice is a guess. Even so, King and Yuille found that children often made false identifications when given verbal instructions that the target person might not be among the choices comprising the lineup. Goodman et al. (1991), however, were successful in decreasing children's errors in response to target-absent lineups. Specifically, 3- to 4-year-old and 5- to 7-year-old children who received a routine inoculation by a nurse were later asked to identify the nurse from a target-absent photo lineup. Verbal instructions to children that the lineup might not include the target were reinforced with explicit training; before attempting identification of the nurse, some children were given practice with target-present and target-absent lineups. For example, children were asked to identify their own mothers from a target-absent photo lineup of women who were ethnically similar to their mother. Incorrect identifications were corrected. Following the practice, children were presented with a target-absent lineup to identify the nurse who had given them the shots. Children who received training were significantly less likely to incorrectly choose a picture (51%) than were the control children who did not receive training (80%). Thus children's false identifications were reduced with instructions and training; however, even with training, children, particularly young children, still made a substantial number of false identifications in response to target-absent lineups.

Finally, although research indicates that children fare better when identifying targets from target-present rather than target-absent lineups, important age differences in accuracy suggest that there is still room for improvement. Goodman and Reed (1986) found that 3-year-olds were inferior to 6-year-olds and adults on a target-present photo identification task. In a study by Peters (1991), 7- to 8-year-olds were more accurate than 3- to 6-year-olds on a target-present task. King and Yuille (1986) also report that 6-year-olds were significantly less likely than older children to correctly identify a target. These results illustrate a developmental trend in children's ability to respond accurately to lineups and underscore the need to develop techniques to aid young children on such tasks.

Summary

Techniques currently used to obtain person identification information from adults may meet with little success when used with children. In response to free recall questions about appearance, children provide little information. This may mean that, in many cases, if investigators have no other relevant information about a suspect's identity, children may not be able to provide enough initial information to guide authorities to a suspect so a lineup can be established. Even if a lineup were to be constructed, children, particularly those younger than 6 years old, are not very accurate at identifying persons from lineups, particularly target-absent lineups.

Developing Age-Appropriate Methods of Obtaining Person Information

Fortunately, research is beginning to illustrate ways in which children's performance can be improved on photo lineup tasks (e.g., Goodman et al., 1991). But there still is a need to develop methods to aid children's retrieval of information critical to the initial identification of suspects preceding lineup construction. It was our goal to develop a person identification task appropriate for use with children. In so doing, we considered existing literature on children's use of external cues in remembering past events and previous research on improving children's ability to use lineups.

We reasoned that if children's retrieval of information about past events could be enhanced with recognition-linked questioning, their accuracy for person characteristics and facial features also could be improved with a recognition technique specifically related to person features. Research has found that children, more than adults, seem to process facial information through individual feature extraction (Carey & Diamond, 1977; Diamond & Carey, 1977). Thus we expected that it would be beneficial for children to begin the person identification process by recognizing individual features rather than the whole person, which has been the focus of past lineup research.

Flin and colleagues (1989) incorporated a feature cue approach in their investigations but found no benefits from children's use of identification kits to recall strangers' appearance. We believe that the children's lack of improvement probably was attributable to the inordinate complexity of the Identi-Kits used rather than to the rationale of using a task emphasizing individual feature recognition. Thus we sought to create specific feature cues that were age appropriate. In so doing, we extended the work of Rudy and Goodman (1991), Tobey and Goodman (1992), and Saywitz et al. (1991) that focused on aiding children in reporting the age and height of a stranger. Specifically, in the first two of these previous studies, children were given a lineup of photographs of persons representing different age ranges. Children gave more correct responses when picking a photograph than they did when answering a free recall question, "How old was the person?" However, this increase in correct responses was also accompanied by an increase in errors; children who gave no response or said "don't know" in response to the free recall question often chose an incorrect photograph. In the Saywitz et al. (1991) study, a series of markers placed on a wall was used to indicate height increments. Although 7-year-olds outperformed 5-year-olds, performance on the height question was poor regardless of whether the children answered a free recall question about height or indicated height using the wall markers. In two experiments, we extended these studies by examining children's ability to report a number of characteristics important in person identification, developing techniques to aid children in recalling eye color, hair color and texture, weight, and race.

EXPERIMENT 1

In Experiment 1, children interacted with a stranger and were then interviewed in detail about the stranger's appearance. Children were first asked to recall specifics about the stranger's features (e.g., age, race, hair color, eye color) and then were presented with a series of feature lineups composed of common variations of each feature. To help children understand the task, we gave children practice using the lineups before they used them to answer questions about the target person.

Based on past findings that children's free recall of information such as age and weight is fairly poor (Brigham et al., 1986) but that children are able to provide more accurate information about past events when interviewed with cues and practice (Goodman et al., 1991), we expected our feature lineup procedure to aid children's reports relative to free recall. We expected that although both older and younger children would benefit from the feature lineup method, older children would for the most part perform better than younger children, as is generally true in memory tasks (e.g., Kail, 1990).

Method

Subjects. A total of 24 3- to 4-year-olds (M = 50 months, range = 43-59 months) and 24 5- to 6-year-olds (M = 72 months, range = 62-75 months) participated in the study. There were equal numbers of boys and girls in each age group. Children were from urban and suburban middle and upper middle socioeconomic status families. Each child received $10 and a toy for participation.

Materials. Separate lineups were specially constructed for the following person features: age, hair color, hair length, hair texture, height, skin color, weight, and eye color. Each lineup was composed of separate representations of typical instances of each feature. Similar to lineups used by Rudy and Goodman (1991) and Tobey and Goodman (1992), the age lineup consisted of color magazine photographs of one female from each of the following age groups: child, teenager, 20s, 30s-40s, middle age, over 70. The

skin color lineup consisted of magazine photographs of three women of different races: one Caucasian, one Asian, and one African American. Eye color was represented by three simplistic line drawings of eyes that were colored green, brown, and blue. The weight lineup consisted of line drawings of women of three shapes: overweight, very thin, and average. Hair texture was represented by line drawings depicting straight, curly, and wavy hair. The hair length lineup consisted of line drawings of short, medium, and long hair. There were three versions of the hair length lineups, one for each of the three hair texture types. (Thus, if a child initially picked wavy hair when asked about hair texture, he or she would be shown only the wavy set of length drawings when asked about hair length.) The hair color lineup consisted of five swatches of synthetic hair used by beauticians for artificial color demonstrations. Basic shades of blonde, brown, black, red, and gray hair were represented. Finally, the height lineup was composed of colored, 1-inch-thick horizontal lines placed on a wall in 6-inch increments starting 3 feet from floor level and ending at 6 feet 6 inches (similar to the technique used by Saywitz et al., 1991).

Our results indicated that use of most of the lineups increased correct responses and decreased "don't know" responses when children were questioned about a stranger's appearance; however, this result was at the unacceptable expense of increasing commission errors.

Ten adults made preratings of each feature representation in each lineup to ensure that there was a consensus regarding the representations. (Preratings were not conducted for the height lineup given the inherent objective measurements for each item.) No less than 80% of the raters (M = .96) provided the intended responses for hair, race, weight, and eye color representations. For the age lineup, raters independently estimated the age of each person depicted in the lineup photographs. The mean age rating for the photograph picturing the child was 11 years (range = 8-13); the teenager, 15 years (range = 9-21); the woman in her 20s, 26 years (range = 19-32); the woman in her 30s-40s, 45 years (range = 37-50); the middle-aged woman, 65 years (range = 47-82); and the oldest woman, 85 years (range = 70-100). On the basis of these preratings, all lineup representations were considered acceptable.

Procedure. Each child was brought individually to our laboratory by a parent or both parents. After fully informed parental consent was obtained, each child was taken to a designated playroom and introduced to an unfamiliar female research assistant (henceforth referred to as the "stranger"). Four different women played the role of the stranger in the study, and equal numbers of children of each age interacted with each stranger. During a play session of approximately 25 minutes, the stranger involved the child in engaging play with a giant inflatable Gumby® doll, a ball, Play-Doh®, puppets, crayons, and the board game Memory® (similar to Concentration®). Each child was unaware that questioning about the stranger would follow. Next, the child and the mother interacted for approximately 10 minutes in another room in play activities. The mother and child did not discuss the appearance of the stranger.

Following this break, the child was introduced to a new female research assistant who conducted an interview session. The assistant told the child, "You know, I don't know the person you played with when you first got here. I'd like to know what she looked like. I want you to help me." For each of the eight target features, the child was first asked to recall the feature (e.g., "How old was the person you played with?"). Immediately after the child's response, the assistant presented the child with the corresponding feature lineup (representations within each lineup were displayed simultaneously). To make sure the children understood the lineup tasks, all children were given practice with each lineup before being asked to use the lineup to report details about the stranger. Practice consisted of asking the children first to use the lineup to describe that feature of themselves and then to describe that feature of the interviewer. For example, for the hair color lineup, each child was asked to point to the swatch of hair that was most similar in color to his or her own and then to the swatch most like the interviewer's hair color. Mistakes were corrected. Finally, each child was asked to point to the swatch that best represented the stranger's hair color. The child was given no feedback about the accuracy of this choice.

When interviews were completed, children were reunited with their mothers, thanked for their help, praised for their performance, debriefed, and given toys and payment for participating.

Results and Discussion

Children's responses to questions about the stranger were scored as correct, don't know, incorrect (commission error), or ambiguous. A child's response to recall questions and feature lineup questions was considered correct if it matched the prerating response given by a consensus of 10 adults who made prejudgments about each stranger's features using both tasks (i.e., free recall and lineup). For example, if the stranger's hair color was brown and the child said "brown" or chose the brown hair from the lineup, the child's response was scored as correct. However, if the stranger's age was 30 and the child said "in her twenties" or chose the picture of a woman in her 20s, that response was scored as a commission error. Finally, if a child stated that the stranger was "tall" in height or "an adult" in age, these responses were scored as ambiguous.

To determine the extent to which children's reports of the stranger's appearance were aided by the use of feature lineups, we compared children's identification accuracy in response to each recall question to their accuracy in response to each feature lineup. Separate 2 (child gender) × 2 (age group: 3- to 4-year-olds vs. 5- to 6-year-olds) × 2 (task: recall vs. feature lineup recall) multivariate analyses of variance (MANOVAs) were conducted with task varied within subjects. The eight feature types—age, hair color, hair length, hair texture, skin color, height, weight, and eye color—served as dependent variables. The responses—correct, don't know, incorrect, and ambiguous—were analyzed separately.[2]

No significant main effects or interactions involving gender emerged for any analysis, all multivariate Fs(8, 37) ≤ .87. Significant effects of task and age are discussed in the following subsections.

Effects of task. Overall, children gave significantly more correct answers in response to the lineup task (M = .52) than to the uncued recall task (M = .34), multivariate $F(8, 37) = 9.92$, $p < .01$ (see Table 1 for all means). Univariate tests revealed significant increases in accuracy for the children's reports of age, height, eye color, and hair color, all Fs(1, 44) ≥ 7.97, ps < .01. The rate of correct responses to queries about other features did not differ significantly as a function of task type.

In addition to increasing correct responses, use of the lineup, in contrast to the recall task, significantly decreased the children's overall "don't know" responses, multivariate $F(8, 37) = 28.80$, $p < .01$ (Table 1). Univariate tests revealed significant decreases for age, height, eye color, hair length, and hair texture, all Fs(1, 44) ≥ 4.40, ps < .05 (Table 1). Unfortunately, however, this decreased tendency to respond "don't know" led not only to a higher number of correct responses but also to an increase in commission errors. Analyses revealed that children gave significantly more incorrect answers overall in response to the lineup task as compared to the recall task, multivariate $F(8, 37) = 5.51$, $p < .01$. Univariate tests indicated that the increase was significant for age, height, eye color, and weight lineups, all Fs(1, 44) ≥ 5.24, ps < .05.

TABLE 1: Proportions of Correct, "Don't Know," and Incorrect Responses as a Function of Task and Feature Type: Experiment 1

	Recall Task			*Lineup Task*		
Feature	*Correct*	*Don't Know*	*Incorrect*	*Correct*	*Don't Know*	*Incorrect*
Age	.00	.79	.13 (.17)	.40	.00	.58 (.58)
Height	.08	.27	.13 (.60)	.56	.02	.42 (.42)
Eye color	.25	.48	.27 (.27)	.46	.04	.48 (.48)
Hair color	.35	.10	.52 (.54)	.63	.02	.35 (.35)
Hair length	.35	.15	.40 (.50)	.48	.00	.52 (.52)
Hair texture	.40	.08	.52 (.52)	.38	.00	.60 (.60)
Skin color	.58	.13	.23 (.29)	.67	.02	.31 (.31)
Weight	.73	.06	.21 (.21)	.58	.02	.40 (.40)
Marginals	.34	.26	.30 (.39)	.52	.02	.46 (.46)

NOTE: Numbers in parentheses represent the proportions of incorrect plus ambiguous responses.

Children rarely gave unscorable or ambiguous responses, with the exception that they appeared to have particular difficulty verbalizing scoreable responses to the height recall question. For example, the children would give answers such as "big as 18, I would say" or "taller than the camera" (which was on a 4-foot tripod), or they would simply point somewhere above their own heads. Although such ambiguous answers are not technically wrong (i.e., the stranger *was* taller than the camera tripod), they would not be very helpful in an actual investigation. Thus we performed a separate MANOVA on incorrect plus ambiguous responses. The overall pattern of results was the same as that already reported for incorrect responses except that the univariate comparison of responses to height as a function of task was no longer significant. In fact, there was a nonsignificant trend in the opposite direction: Errors plus ambiguous responses decreased in response to the height lineup as compared to those in the uncued recall, $F(1, 44) = 3.04$, $p = .09$ (Table 1).

Age differences. As expected, younger children provided significantly fewer correct responses and made significantly more errors overall than did older children, all multivariate Fs(8, 37) ≥ 2.37, ps < .05 (see Table 2). Univariate tests revealed that for both correct and incorrect responses, age differences were

TABLE 2: Proportions of Correct, "Don't Know," and Incorrect Responses as a Function of Age and Feature Type: Experiment 1

	3- to 4-Year-Olds			5- to 6-Year-Olds		
Feature	Correct	Don't Know	Incorrect	Correct	Don't Know	Incorrect
Age	.17	.35	.42 (.46)	.23	.44	.29 (.29)
Height	.21	.08	.40 (.69)	.44	.21	.15 (.33)
Eye color	.31	.25	.44 (.44)	.40	.27	.31 (.31)
Hair color	.40	.06	.52 (.54)	.58	.06	.35 (.35)
Hair length	.44	.04	.44 (.52)	.40	.10	.48 (.50)
Hair texture	.38	.04	.56 (.56)	.40	.04	.56 (.56)
Skin color	.44	.10	.46 (.46)	.81	.04	.08 (.15)
Weight	.60	.04	.35 (.35)	.71	.04	.25 (.25)
Marginals	.37	.12	.45 (.50)	.50	.15	.31 (.34)

NOTE: Numbers in parentheses represent the proportions of incorrect plus ambiguous responses.

significant for recall of height and skin color, $Fs(1, 44) \geq 7.69$, $ps \leq .01$. No significant main effects or interactions involving age emerged for "don't know" responses, all $Fs(8, 37) \leq 1.68$.

Summary. Our results indicated that use of most of the lineups increased correct responses and decreased "don't know" responses when children were questioned about a stranger's appearance; however, this result was at the unacceptable expense of increasing commission errors. Further, when developmental differences occurred, older children outperformed younger children. Results clearly indicated the need for additional research, which prompted Experiment 2.

EXPERIMENT 2

We reasoned that the increase in incorrect responses in Experiment 1 was caused by the demand characteristics of our lineup tasks; that is, children probably thought they were expected to pick something from the lineups even if they were unsure. Although "I don't know" was an obvious response option when children were asked a recall question about the stranger's appearance, our lineup tasks did not overtly present children with the option of responding "I don't know." That children did not perceive "don't know" to be an option on the lineup task is evident in the relatively few "don't know" responses on that task, a finding consistent with previous research. Specifically, Rudy and Goodman (1991) and Tobey and Goodman (1992) also found an increase in incorrect responses in response to lineup tasks versus recall tasks. They noted that when children responded with silence or "don't know" during recall, they often would then choose an incorrect photo when using the lineup. It is also consistent with results of a study by Dekle, Beal, Elliott, and Huneycutt (1994). Specifically, Dekle et al. found that children who were presented with a target-present lineup were less likely than adults to use a "don't know" or "not sure" option. The authors believe that children were functionally unaware of this option, even when told, because it was not a choice physically presented in the lineup. Dekle et al. further speculated that children may provide incorrect identifications rather than responding "I don't know" because they do not recognize their own uncertainty.

In Experiment 2, we attempted to reduce the inaccurate responses found in Experiment 1 by providing the children with a physical reminder of the "don't know" response option: a card with a question mark drawn on it placed within each feature lineup. We hypothesized that the question mark would reemphasize for the children that a "don't know" response is acceptable when they are unsure of the correct answer.

The overall effect of age that emerged in Experiment 1 implied that our feature lineups are probably less appropriate for very young children (especially 3- to 4-year-olds) than for older children. Saywitz et al. (1991) also found older children to be aided somewhat more than younger children by their height recall technique, and others have found the performance of 3-year-olds to be generally inferior to that of older children in target-present photo identification tasks (Goodman & Reed, 1986; Goodman et al., 1991; Peters, 1991). Given these findings, in Experiment 2 we decided to explore the usefulness of the modified feature lineups with older groups of child participants: 4- to 6-year-olds and 8- to 9-year-olds. Because we expected older children to be quite accurate in remembering person characteristics, we included a 2-week delay between play activity and questioning to avoid possible ceiling effects. The longer delay also adds to the ecological validity of the research because children who must recall the identity of strangers in a legally relevant context often are required to make reports weeks or months after the original event.

Finally, given the significant improvement that Goodman et al. (1991) found when practice was included for a target-absent lineup task, we assumed in Experiment 1 that there would be benefits in providing practice in the use of our features lineups. But providing practice for children proved to be time consuming; in fact, the practice procedures more than doubled the time needed to obtain person identification information from the children. Thus, in Experiment 2, we felt it necessary to systematically investigate the usefulness of our practice procedure. Even considering the lengthy nature of the practice,

in light of the significant benefits found by Goodman et al., we predicted that practice would aid the children in using the lineups.

With these changes, we once again examined children's recall responses versus their lineup task responses. We predicted that use of the new lineups including the question mark option would reduce inaccurate responses while maintaining an increase in accurate responses and that older children would perform better than younger children.

Method

Subjects. A total of 48 children, consisting of 24 4- to 6-year-olds (M = 62 months, range = 58-74 months) and 24 8- to 9-year-olds (M = 101 months, range = 96-108 months) served as subjects.[5] Each age group had an even number of girls and boys, and half of the children in each age and gender group were randomly assigned to the practice condition. Children were from middle and upper middle socioeconomic status urban and suburban families. Each child received $10 and two toys for participation.

Materials. As in Experiment 1, separate lineups were constructed for the following person features: age, hair color, hair length, hair texture, height, skin color, weight, and eye color. The hair color, height, skin color, and weight lineups were constructed as they were in Experiment 1. The age representations were similar, but we attempted to pick more realistic pictures and we added an additional age category to make our ranges more specific (one category for 30s and one for 40s-50s instead of just one for 30s-40s). For the eye color, hair texture, and hair length lineups, more realistic magazine pictures were used instead of line drawings. Finally, in addition to the lineup options representing each feature choice, a card with a question mark was included as an option in every lineup.

Fifteen adults made preratings of all lineup representations except height. Percentage agreement among raters for hair, race, weight, and eye color was no less than 93% for any single representation (M = .99). The mean age rating for the child photograph was 7 years (range = 6-10); the teenager, 15 years (range = 13-17); the woman in her 20s, 25 years (range = 19-29); the woman in her 30s, 32 years (range = 29-34); the woman in her 40s-50s, 52 years (range = 45-59); the middle-aged woman, 61 years (range = 58-64); the oldest woman, 80 years (range = 76-89). These preratings indicated that our representations were valid.

Procedure. Each child was scheduled to come to the laboratory for two sessions. During the first session, as in Experiment 1, parental consent was obtained and the child was taken to a playroom and introduced to a female stranger. An equal number of children interacted with each of five different women who played the role of the stranger in the study. The play session lasted approximately 18 minutes, during which the child and the stranger engaged in play with a Gumby doll, Play-Doh, Mr. Potato Head®, crayons, a game of Simon Says, and a puzzle. The child was unaware that questioning about the stranger would follow 2 weeks later.

The remainder of the procedure was identical to that of Experiment 1 except that (a) the interview session occurred after approximately a 2-week delay (M = 14 days, range = 12-16 days) and (b) half of the children did not receive practice before using the feature lineups.

TABLE 3: Proportion of Correct, "Don't Know," and Incorrect Responses as a Function of Task and Feature Type: Experiment 2

	Recall Task			*Lineup Task*		
Feature	*Correct*	*Don't Know*	*Incorrect*	*Correct*	*Don't Know*	*Incorrect*
Age	.06	.58	.35 (.35)	.27	.40	.33 (.33)
Height	.08	.42	.23 (.50)	.35	.06	.58 (.58)
Eye color	.21	.40	.40 (.40)	.19	.42	.40 (.40)
Hair color	.44	.15	.38 (.42)	.44	.23	.33 (.33)
Hair length	.46	.21	.21 (.33)	.56	.19	.25 (.25)
Hair texture	.40	.10	.48 (.50)	.35	.33	.31 (.31)
Skin color	.67	.04	.25 (.29)	.56	.21	.23 (.23)
Weight	.48	.06	.46 (.46)	.44	.19	.38 (.38)
Marginals	.35	.25	.35 (.41)	.40	.25	.35 (.35)

NOTE: Numbers in parentheses represent the proportions of incorrect plus ambiguous responses.

Results and Discussion

Preliminary analyses revealed no significant main effects or interactions attributable to practice, and data were collapsed across this factor. Remaining analyses were identical to those in Experiment 1: 2 (gender) × 2 (age group: 4- to 6-year-olds vs. 8- to 9-year-olds) × 2 (task: recall vs. feature lineup) MANOVAs.

Effects of task. Children gave significantly more correct answers in response to the lineup task than they did in response to the recall task, multivariate $F(8, 37)$ = 3.43, $p < .01$ (see Table 3 for all means). Specifically, univariate tests revealed significant increases in accuracy for children's reports of age, $F(1, 44)$ = 13.10, $p > .001$, and height, $F(1, 44)$ = 13.37, $p < .001$. There was also a significant multivariate effect of task on "don't know" responses, $F(8, 37)$ = 4.09, $p < .01$ (for the

TABLE 4: Proportions of Correct, "Don't Know," and Incorrect Responses as a Function of Age and Feature Type: Experiment 2

	4- to 6-Year-Olds			8- to 9-Year-Olds		
Feature	Correct	Don't Know	Incorrect	Correct	Don't Know	Incorrect
Age	.04	.67	.29 (.29)	.29	.31	.40 (.40)
Height	.17	.31	.37 (.52)	.27	.17	.44 (.56)
Eye color	.25	.48	.27 (.27)	.15	.33	.52 (.52)
Hair color	.38	.25	.35 (.38)	.50	.13	.35 (.38)
Hair length	.42	.29	.23 (.29)	.60	.10	.23 (.29)
Hair texture	.31	.35	.31 (.33)	.44	.08	.48 (.48)
Skin color	.58	.15	.25 (.27)	.65	.10	.23 (.25)
Weight	.35	.19	.46 (.46)	.56	.06	.38 (.38)
Marginals	.31	.34	.32 (.35)	.43	.16	.38 (.41)

NOTE: Numbers in parentheses represent the proportions of incorrect plus ambiguous responses.

lineup task, a "don't know" response was a choice of the card with the question mark). Univariate tests revealed significant differences for age, hair texture, height, skin color, and weight, all Fs(1, 44) > 4.40, p < .05. These effects were of a mixed nature; children's "don't know" responses decreased from free recall to lineup questions for age and height but increased for hair texture, skin color, and weight.

Children's incorrect responses differed significantly across type of task, multivariate $F(8, 37) = 2.51$, $p < .05$ (Table 3). Specifically, children's incorrect responses decreased significantly on the lineup task when responding about the stranger's hair texture compared to their recall responses, univariate $F(1, 44) = 6.18$, $p < .05$. But errors increased when children used the lineup to respond about the stranger's height, univariate $F(1, 44) = 12.47$, $p < .01$. An additional MANOVA was conducted on children's incorrect responses combined with their ambiguous or unhelpful responses. There were no significant main effects or interactions, all Fs(8, 37) < 1.95.

Age differences. Interestingly, age differences emerged only for correct responses, multivariate $F(8, 37) = 2.16$, $p = .05$, and then only for children's responses about the stranger's age, univariate $F(1, 44) = 11.00$, $p < .01$ (see Table 4). Older children provided significantly more correct responses about the stranger's age than did younger children. No additional significant main effects or interactions associated with age emerged for incorrect or "don't know" responses, all Fs(8, 37) ≤ 2.01.

Gender differences. Gender differences emerged for incorrect responses, multivariate $F(8, 37) = 2.55$, $p < .05$. Univariate tests revealed that males provided significantly more incorrect responses ($M = .46$) than did females ($M = .23$) about the stranger's age, $F(1, 44) = 4.81$, $p < .05$. Males also made significantly more incorrect responses ($M = .52$) than did females ($M = .29$) about the stranger's height, $F(1, 44) = 7.61$, $p < .01$.

In addition, a significant but meaningfully uninterpretable Task Type × Gender × Age interaction emerged for "don't know" responses, $F(8, 37) = 2.55$, $p < .05$.

Summary. In Experiment 2, older children using the modified lineups appeared to benefit most from the technique. Even with the 2-week delay, correct responses increased without the pervasive increase in errors seen previously. Children's ability to provide information, particularly about height and age, improved with the lineups. Even so, our technique increased incorrect responses about hair texture and was largely ineffective for the other person characteristic dimensions. Surprisingly, practice had no significant effect on the children's performance. Finally, we found evidence of gender differences with girls outperforming boys on age and height questions.

DISCUSSION

In this investigation, we developed and tested an interview technique expected to aid children in recalling the appearance of strangers. Previous research has revealed that recall questions do not provide children with sufficient cues to access memories (e.g., Goodman et al., 1991). When children in our study were asked to recall a stranger's features, they were able to report very little accurate information, especially for some characteristics such as age, height, and eye color. But for these same characteristics, children gave significantly more correct responses when using the feature lineups. Why were children able to provide a greater number of accurate responses when using the lineups? The lineups probably acted as cues to make the task more like one of recognition rather than recall, providing easier access to information in the children's memory. In addition, the linguistic skills needed to use the lineup were not as sophisticated or complex as those needed for accurate verbal recall. Even if children could retrieve correct information about strangers' appearance, communicating the information might have been difficult given their limited linguistic abilities (Kail, 1990; Walker, 1993, 1994). Thus the lineups might have served as a memory cue and/or as a communication aid.

Even so, the effects of the lineups were not all positive, particularly in Experiment 1. As Pipe et al.

(1993) note, props or cues that increase accuracy may also have the disadvantage of leading children into commission errors. Children in our first experiment, like those in Tobey and Goodman's (1992) study, made more commission errors when using lineups than they did when responding to free recall questions. The feature lineups used in Experiment 1 did not include an explicit "don't know" option, and although children were instructed that they could verbally respond "I don't know" if they could not remember, they rarely gave this type of response. The fact that children might have felt compelled to choose an item rather than saying "I don't know" is probably responsible for the high incidence of errors.

When the lineup task was modified in Experiment 2 to include the tangible "don't know" option, the frequency of commission errors on the lineup task compared to that on the free recall task increased only for height and in fact decreased for hair texture (although these differences were eliminated when ambiguous responses were included in our analyses). Children apparently were comfortable choosing "I don't know" (i.e., the card with the question mark) when it was presented as an option among the other lineup items. The increase in "don't know" responses appears to have led to both fewer accurate responses and fewer inaccurate responses. Nevertheless, in a forensic context, an increase in omission errors, when accompanied by a decrease in commission errors, is arguably a favorable outcome.

The increase in "don't know" responses appears to have led to both fewer accurate responses and fewer inaccurate responses. Nevertheless, in a forensic context, an increase in omission errors, when accompanied by a decrease in commission errors, is arguably a favorable outcome.

Age differences emerged in both experiments. Older children were more accurate than younger children in recalling person features. In Experiment 1, 5- to 6-year-olds outperformed 3- to 4-year-olds when asked about the stranger's height and skin color. Although young children can be quite accurate on some eyewitness testimony tasks (e.g., Goodman et al., 1990), the developmental differences we found are not surprising given research demonstrating changes in memory development between 3 and 6 years of age (e.g., in children's ability to use mnemonic strategies, in their meta-memory abilities, in their general knowledge base; see Kail, 1990, and Fivush & Hudson, 1990). The developmental difference in the children's ability to report the stranger's skin color is interesting in light of research on the development of children's racial knowledge. By the age of 4 years, children are capable of correctly identifying the skin color of others (Katz, 1976); however, racial group concepts (stereotypes) may not be consolidated and group differences may not be elaborated until after 5 years of age (Katz & Zalk, 1974). These findings indicate that the ability to accurately recall and report a stranger's skin color appears to accompany the development of more sophisticated concepts associated with race.

Interestingly, the primary age difference that emerged in Experiment 2 was for correct responses about the stranger's age. More pervasive age effects were found in Experiment 1, which included younger child participants. Our lineup techniques seem to hold most promise for use with older children. Although this conclusion is tentative until the modified lineups are tested with children as young as 3 years old, a growing body of research indicates that children this young have not yet developed the symbolic representation skills necessary for use of props such as those in our lineup task (DeLoache, 1990, 1995; DeLoache, Kolstad, & Anderson, 1991). In future research, more dramatic age differences might be found with longer delays between the target event (being exposed to a stranger) and recall; research investigating retention and forgetting indicates that although children's memories of events over time can be quite good, young children's performances often decrease over time more than do older children's and adults' performances (Brainerd & Ornstein, 1991).

Finally, although we did not predict the gender differences that emerged, they are not the first to be reported in child testimony research (Rudy & Goodman, 1991; Tobey & Goodman, 1992). Gender differences in memory tasks typically favor girls. As noted by Tobey and Goodman, girls generally perform better at tasks requiring verbal skills (Cox & Waters, 1986) and at memory tasks involving social context (Feldstein, 1976). Our recall task made a fairly complex demand on the children's linguistic abilities, particularly their vocabularies. Further, the memory

tested in the present research concerned person information, a very socially relevant type of information.

FUTURE DIRECTIONS

Although the results from this study are encouraging, they indicate the need for more research on this topic. Additional modifications of the lineup technique are needed to increase children's accuracy while eliminating both commission errors and "don't know" responses. Different pre-lineup practice may be indicated. As illustrated by Goodman et al. (1991) and Geiselman, Saywitz, and Bornstein (1993), training can be effective in aiding children's ability to recall past events. Even so, our practice manipulation had no effect on children's performance. In our studies, children practiced using the lineups on themselves and the interviewer but were then required to use the lineups on someone not present in the room. A more effective practice session might include having children practice using the lineups to recall information about a familiar person not present in the room such as a parent. This might lead to greater transfer of practice because the test requires remembering information about a stranger who is also not present during testing. In addition, this change in practice might increase children's understanding of the lineup technique, allowing them to concentrate less on the nature of the task and more on retrieving the to-be-remembered information; this, in turn, might improve the children's use of the lineups. Further, our training did not include an example for which the children would not know the answer. This training should be included so that children can be taught that choosing the "don't know" option is an acceptable and sometimes accurate response.

An additional future modification of the lineup should be the use of computer-generated feature representations rather than magazine pictures such as those used in the present investigation. This would control for differences between the lineup items that may have distracted the children from concentrating on the feature in question. For example, the eye color lineup consisted of three women with the three distinct eye colors but also three different hair colors, hair textures, and so on. Using computer-generated pictures would eliminate these additional differences between items and allow children to concentrate on only one feature at a time.

Further, Wells (1993) has noted that sequential presentation procedures, in which subjects respond to lineup items individually rather than simultaneously, produce fewer false identifications without reducing correct responses in adult subjects. Research is needed to investigate whether this procedural change could reduce children's errors while maintaining the increase in children's correct responses to our technique.

Finally, if our findings are to be most applicable to the forensic contexts that initially prompt our research, future modifications must include efforts to simulate the varied contexts of real crime investigations. Although not all children who must give descriptions of strangers in legally relevant contexts give their reports after long delays or about emotionally troubling events, some certainly do. Thus research designs incorporating longer delays, emotionally significant interactions with strangers, and so on are indicated.

CONCLUSION

Obtaining enough information to determine the identity of perpetrators of crimes is paramount in actual legal investigations. Children's uncued verbal reports are unlikely to contain sufficient information to lead to the identification of strangers, yet techniques now used by authorities with adults for cuing descriptions of suspects are not appropriate for use with young children. In our experiments, we demonstrated the potential benefits and pitfalls of a feature lineup approach. Future research needs to focus on refining the technique to make it one that will decrease the error rate that appears to result from demand characteristics inherently a part of our interview technique. As we continue our research in this area, we are confident that future investigations will culminate in the development of a lineup interview technique that will effectively facilitate children's person identification recall and will provide legal authorities with a way to optimize children's "person testimony."

In conclusion, the legal system expects a witness to provide complete and accurate reports when he or she is questioned about a witnessed event. This expectation is based on the notion that all witnesses possess the memory skills needed to access the correct information. Research has demonstrated numerous times, however, that uncued questioning typically does not lead to sufficient information when young children are involved. Therefore, the legal system should not simply expect a child witness to provide accurate and complete reports but should also provide the means of successfully accessing those accurate and complete reports. It is up to empirical researchers to help ensure the attainment of this goal so that children's actual abilities are not underestimated, so that necessary information is obtained, and, ultimately, so that questions of guilt and innocence are answered. The present research is the first of many steps needed to

help provide the legal system with age-appropriate questioning techniques.

NOTES

1. Investigators may also use a "showup," during which only one suspect is presented to a witness. Although the courts historically have considered showups to be inherently more suggestive than lineups, recent research finds that false identification errors may be reduced for both adults and children using showups compared to those using lineups (Dekle et al., 1994; Gonzalez, Ellsworth, & Pembroke, 1993).

2. When proportions are reported as cell means, the analyses were based on dichotomous (correct vs. incorrect) variables, not on proportions. (However, the mean of a 0-1 variable is equal to the proportion of 1s.) Analysis of variance can be validly conducted on dichotomous data (e.g., see Lunney, 1970).

3. Data from two additional subjects were dropped due to experimenter error.

REFERENCES

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*, 54-75.

Brainerd, C., & Ornstein, P. (1991). Children's testimony: The developmental backdrop. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 10-20). Washington, DC: American Psychological Association.

Brigham, J. C., Van Verst, M., & Bothwell, R. K. (1986). Accuracy of children's eyewitness identifications in a field setting. *Basic and Applied Social Psychology, 7*, 295-306.

Brown, A. L. (1979). Theories of memory and problems of development: Activity, growth, and knowledge. In L. Cermak & F.I.M. Craik (Eds.), *Levels of processing in memory* (pp. 225-258). Hillside, NJ: Lawrence Erlbaum.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representations of faces. *Science, 195*, 312-314.

Chance, J. E., & Goldstein, A. G. (1984). Face recognition memory: Implications for children's eyewitness testimony. *Journal of Social Issues, 40*, 69-86.

Comish, S. E. (1987). Recognition of facial stimuli following an interview task involving the Identi-Kit. *Journal of Applied Psychology, 72*, 488-491.

Cox, D., & Waters, H. S. (1986). Sex differences in the use of organization strategies: A developmental analysis. *Journal of Experimental Child Psychology, 41*, 18-37.

Davies, G. (1981). Face recall systems. In G. Davies, H. Ellis, & J. Sheperd (Eds.), *Perceiving and remembering faces* (pp. 227-250). London: Academic Press.

Davies, G., Tarrant, A., & Flin, R. (1989). Close encounters of the witness kind: Children's memory for a simulated health inspection. *British Journal of Psychology, 80*, 415-429.

Dekle, D. J., Beal, C. R., Elliott, R., & Huneycutt, D. (1994). *Children as witnesses: A comparison of lineup versus showup identification methods.* Unpublished manuscript, Department of Psychology, Dartmouth College.

DeLoache, J. S. (1990). Young children's understanding of models. In R. Fivush & J. A. Hudson (Eds.), *Knowing and remembering in young children* (pp. 94-126). Cambridge, England: Cambridge University Press.

DeLoache, J. S. (1995). The use of dolls in interviewing young children. In M. S. Zaragoza, J. R. Graham, G.C.N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 160-178). Thousand Oaks, CA: Sage.

DeLoache, J. S., Kolstad, V., & Anderson, K. N. (1991). Physical similarity and young children's understanding of scale models. *Child Development, 62*, 111-126.

Dent, H. R. (1991). Experimental studies of interviewing child witnesses. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 138-146). Washington, DC: American Psychological Association.

Diamond, R., & Carey, S. (1977). Developmental changes in the representation of faces. *Journal of Experimental Child Psychology, 23*, 1-22.

Feldstein, J. H. (1976). Sex differences in social memory among preschool children. *Sex Roles, 2*, 73-79.

Finkelhor, D., Hotaling, G., Lewis, I. A., & Smith, C. (1990). Sexual abuse in a national survey of adult men and women: Relevance, characteristics, and risk factors. *Child Abuse and Neglect, 14*, 385-401.

Fivush, R., & Hudson, J. A. (Eds.). (1990). *Knowing and remembering in young children.* Cambridge, England: Cambridge University Press.

Flin, R., Markham, R., & Davies, G. (1989). Developmental trends in the construction and recognition of photofit face composites. *Journal of Applied Developmental Psychology, 10*, 131-145.

Flin, R., & Shepherd, J. W. (1986). Tall stories: Eyewitnesses' ability to estimate height and weight characteristics. *Human Learning, 5*, 29-38.

Geiselman, R. E., Saywitz, K. J., & Bornstein, G. K. (1993). Effects of cognitive questioning techniques on children's recall performance. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses* (pp. 71-94). New York: Guilford.

Gonzalez, R., Ellsworth, P. C., & Pembroke, M. (1993). Response biases in lineups and showups. *Journal of Personality and Social Psychology, 64*, 525-537.

Goodman, G. S. (1984). Children's testimony: A historical perspective. *Journal of Social Issues, 40*, 9-32.

Goodman, G. S., & Aman, C. (1990). Children's use of anatomically detailed dolls to recount an event. *Child Development, 61*, 1859-1871.

Goodman, G. S., Bottoms, B. L., Schwartz-Kenney, B. M., & Rudy, L. (1991). Children's testimony about a stressful event: Improving children's reports. *Journal of Narrative and Life History, 1*, 69-99.

Goodman, G. S., & Reed, R. S. (1986). Age differences in eyewitness testimony. *Law and Human Behavior, 10*, 317-332.

Goodman, G. S., Rudy, L., Bottoms, B. L., & Aman, C. (1990). Children's concerns and memory: Ecological issues in the study of children's eyewitness testimony. In R. Fivush & J. Hudson (Eds.), *Knowing and remembering in young children* (pp. 249-284). Cambridge, England: Cambridge University Press.

Jones, D.P.H., & Krugman, R. D. (1986). Can a three-year-old child bear witness to her sexual assault and attempted murder? *Child Abuse and Neglect, 10*, 253-258.

Kail, R. (1990). *The development of memory in children.* New York: Freeman.

Katz, P. A. (1976). The acquisition of racial attitudes in children. In P. A. Katz (Ed.), *Towards the elimination of racism* (pp. 125-154). New York: Pergamon.

Katz, P. A., & Zalk, S. R. (1974). Doll preferences: An index of racial attitudes? *Journal of Educational Psychology, 66*, 663-668.

King, M. A., & Yuille, J. C. (1986). *An investigation of children's eyewitness abilities.* Unpublished manuscript, Department of Psychology, University of British Columbia.

Kobasigawa, A. (1974). Utilization of retrieval cues by children in recall. *Child Development, 45*, 127-134.

Laughery, K., & Fowler, R. (1980). Sketch artist and Identi-Kit procedures for recalling faces. *Journal of Applied Psychology, 65*, 307-316.

Loftus, E. F. (1979). *Eyewitness testimony.* Cambridge, MA: Harvard University Press.

Lunney, G. (1970). Using analysis of variance with a dichotomous dependent variable: An empirical study. *Journal of Educational Measurement, 7*, 263-269.

Mandler, J. M. (1990). Recall and its verbal expression. In R. Fivush & J. A. Hudson (Eds.), *Knowing and remembering in young children* (pp. 317-330). Cambridge: Cambridge University Press.

Meyer, G. E., & Hilterbrand, K. (1984). Does it pay to be "bashful"? The seven dwarfs and long-term memory. *American Journal of Psychology, 97*, 47-55.

Parker, J. F., & Carranza, L. E. (1989). Eyewitness testimony of children in target-present and target-absent lineups. *Law and Human Behavior, 13,* 133-149.

Peters, D. (1991). The influence of stress and arousal on the child witness. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 60-82). Washington, DC: American Psychological Association.

Piaget, J., & Inhelder, B. (1973). *Memory and intelligence.* London: Routledge & Kegan Paul.

Pipe, M.-E., Gee, S., & Wilson, M. (1993). Cues, props and context: Do they facilitate children's event reports? In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses* (pp. 25-45). New York: Guilford.

Price, D.W.W., & Goodman, G. S. (1990). Visiting the wizard: Children's memory for a recurring event. *Child Development, 61,* 664-680.

Rudy, L., & Goodman, G. S. (1991). Effects of participation on children's reports: Implications for children's testimony. *Developmental Psychology, 27,* 527-538.

Saywitz, K., Goodman, G. S., Nicholas, E., & Moan, S. (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., & Snyder, L. (1993). Improving children's testimony with preparation. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses* (pp. 117-146). New York: Guilford.

Tobey, A., & Goodman, G. S. (1992). Children's eyewitness memory: Effects of participation and forensic context. *Child Abuse and Neglect, 16,* 779-796.

Tulving, E. (1983). *Elements of episodic memory.* Oxford, England: Oxford University Press.

Walker, A. G. (1993). Questioning young children in court: A linguistic case study. *Law and Human Behavior, 17,* 59-81.

Walker, A. G. (1994). *Handbook on questioning children: A linguistic perspective.* Washington, DC: American Bar Association, Center on Children and the Law.

Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist, 48,* 553-571.

Wells, G. L., & Seelau, E. P. (1995). Eyewitness identification: Psychological research and legal policy on lineups. *Psychology, Public Policy, and Law, 1,* 765-791.

Whitehurst, G. J. (1982). Language development. In B. B. Wolman & G. Stricker (Eds.), *Handbook of developmental psychology* (pp. 367-386). Englewood Cliffs, NJ: Prentice-Hall.

Beth M. Schwartz-Kenney, Ph.D., Department of Psychology, Randolph-Macon Woman's College; Bette L. Bottoms, Ph.D., Department of Psychology, University of Illinois at Chicago; Gail S. Goodman, Ph.D., University of California, Davis. *This research was funded in part by grants to Beth M. Schwartz-Kenney from the Mednick Fellowships Committee of the Virginia Foundation of Independent Colleges and from the Professional Development Committee at Randolph-Macon Woman's College as well as to Gail S. Goodman from the U.S. Department of Health and Human Services. Thanks are extended to Mary Elizabeth Wilson, Cathleen Carter, Toby Sachsenmaier, Sherry F. Thomas, Lisa Artino, Mary Golden-Considine, Jeannie Hagan, Kimberley Duff, Sonya Martin, and Jeannine Vogel for research assistance. We also thank Dennis Goff for statistical consultation.* *Address correspondence to Beth M. Schwartz-Kenney, Department of Psychology, Randolph-Macon Woman's College, Lynchburg, VA 24503.*