Lineup member similarity effects on children's eyewitness identification

Ryan J. Fitzgerald, Brittany F. Whiting, Natalie M. Therrien, & Heather L. Price

University of Regina

This is the peer reviewed version of the following article: Fitzgerald, R. J., Whiting, B. F., Therrien, N. M., & Price, H. L. (2014). Lineup member similarity effects on children's eyewitness identification. *Applied Cognitive Psychology*, *28,* 409-418, which has been published in final form at http://onlinelibrary.wiley.com/doi/10.1002/acp.3012/full. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

The authors thank Adam Donauer, Jordan Dean, Caitlin Hunter, Shelby Siroski, Tansi Summerfield, Kirstie Walker, Jessica Williams, Josh Gonzales, Nikolina Vracar, Rachelle Jeworski, Sarah Reiser, Sarah Sangster, Cori Carey, Jill Price, Kirsten Gullickson, Megan Adams Lebell, and Kelcie Novak for their assistance with data collection. The authors also sincerely thank all those involved with the Educating Youth in Engineering & Science (EYES) camp, including the children, parents, counselors, and administrators who participated in or supported this research.

This research was supported by an Alexander Graham Bell Canada Graduate Scholarship (Doctoral) from the Natural Sciences and Engineering Research Council of Canada to the first author, Joseph-Armand Bombardier Canada Graduate Scholarships (Masters) from the Social Sciences and Humanities Research Council to the second and third authors, and a grant from the Natural Sciences and Engineering Research Council of Canada and a Canada Foundation for Innovation Leader's Opportunity Fund to the fourth author.

Correspondence concerning this article should be addressed to Ryan J. Fitzgerald, who is now at Department of Psychology, University of Portsmouth, King Henry Building, King Henry 1 Street, Portsmouth, United Kingdom, PO1 2DY. E-mail: ryan.fitzgerald@port.ac.uk

Abstract

To date, research investigating the similarity among lineup members has focused on adult eyewitnesses. In the present research, children made identifications from lineups containing members of lower or higher similarity to a target person. In Experiment 1, following a live interaction, children's (6–14 years) correct identification rate was reduced in higher-similarity relative to lower-similarity lineups. In Experiment 2, children (6–12 years) and adults watched a video containing a target person. Again, higher-similarity lineup members reduced children's correct identifications; however, similarity had no effect on adults' correct identification rate. Although children benefited from lower-similarity lineups when the target was present, lower-similarity lineups generally increased misidentifications of an innocent suspect when the target was absent. Thus, increasing similarity in lineups for children had a cost on target-present lineups and a benefit on target-absent lineups.

Lineup member similarity effects on children's eyewitness identification

In an effort to curb child witnesses' increased propensity to falsely identify an innocent lineup member, researchers have developed novel lineup procedures to make the identification task more appropriate for children. For example, the elimination procedure involves encouraging children to first select the lineup member who most resembles the culprit and then to decide whether that person is in fact the culprit (Pozzulo & Lindsay, 1999). Another procedure involves placing a blank silhouette figure among the lineup members to provide children with a salient rejection option (Zajac & Karageorge, 2009). Although these procedures have been largely successful in reducing children's tendency to choose, another method of protecting innocent suspects from false identification that has been effective with adults has yet to be explored with child witnesses.

The obvious approach to reducing innocent suspect misidentifications is to increase the rate at which eyewitnesses reject lineups that do not contain the culprit; however, innocent suspect misidentifications can also be reduced by shifting identifications from the suspect to the fillers (Wells, Memon, & Penrod, 2006). Police investigators do not know whether the suspect in the lineup is guilty or innocent. As a consequence, a misidentified innocent suspect will be the subject of further investigation and has the potential to be wrongfully convicted. By contrast, filler identifications do not pose the risk of wrongful conviction because lineup fillers are (or at least should be) known innocents (Wells & Turtle, 1986). Thus, a filler selection can have the same exonerating effect as a lineup rejection.

In the section that follows, we outline research on adults showing that the composition of a lineup, particularly the similarity of the lineup members, can have a dramatic effect on whether witnesses misidentify a suspect or a filler. We then demonstrate that in spite of the considerable number of child witness studies that have investigated how lineup members should be presented, no research has directly examined which lineup members should be presented to children. We subsequently report empirical data demonstrating that lineup similarity has an effect on identification responses and that children are particularly sensitive to these effects. These findings lead us to conclude that filler selection strategies that are appropriate for adult witnesses need to be adapted for child witnesses.

Lineup Similarity Effects with Adult Witnesses

A panel of eyewitness experts developed guidelines specifying the extent to which fillers should resemble the person suspected of committing the crime (Technical Working Group for Eyewitness Evidence, 1999). In their report, the authors discourage constructing lineups with dissimilar fillers in order to prevent the suspect from ‘standing out’. Regardless of whether the suspect is guilty or innocent, witnesses tend to choose lineup members who stand out (Lindsay & Wells, 1980). The authors further discourage selecting fillers who closely resemble the suspect on the premise that highly similar fillers would make correct identifications of the culprit too difficult. Thus, police investigators are faced with the challenge of creating lineups that comprise members who are similar, but not too similar.

A proposed method of addressing this challenge is to use the features of the culprit that the witness recalls as filler selection criteria (Luus & Wells, 1991). If the suspect were the only lineup member who matched the description, witnesses would be able to exclude the fillers on account of their mismatch with what was recalled of the culprit’s appearance. Thus, at the very minimum, fillers should possess all of the features mentioned in the witness description. However, Luus and Wells suggested that if fillers resemble the suspect too strongly, important differences among lineup members that are needed for recognition to occur would be lost. Therefore, they suggest permitting lineup members to vary on the features not mentioned in the witness description. Luus and Wells theorized that matching only on the features in the description would protect innocent suspects without hindering culprit identifications.

Comparisons between the match-to-description procedure and the more traditional strategy of matching fillers to the suspect's appearance have produced mixed results. Relative to appearance-matched lineups, description-matched lineups increased correct identifications in one study (Wells, Rydell, & Seelau, 1993), increased innocent suspect misidentifications in another study (Lindsay, Martin, & Webber, 1994), and had no reliable effects on suspect choices in additional studies (Darling, Valentine, & Memon, 2008; Juslin, Olsson, & Winman, 1996; Tunnicliff & Clark, 2000). These conflicting findings, combined with the numerous situations in which matching to a witness description is not practical (Fitzgerald, Price, Oriet, & Charman, 2013), could explain why only a small proportion of police investigators use witness descriptions to choose fillers (Police Executive Research Forum, 2013; Wogalter, Malpass, & McQuiston, 2004).

The majority of police investigators choose fillers on the basis of their match to the suspect's appearance (Police Executive Research Forum, 2013; Wogalter et al., 2004). Thus, it is imperative to determine the optimal level of lineup member similarity. In spite of the concern expressed by some researchers that choosing fillers who closely resemble the suspect's appearance would make culprit identifications too difficult, a recent meta-analysis indicated that correct identifications were unaffected by whether suspect–filler similarity was moderate or high; however, innocent suspect misidentifications increased when similarity was only moderate (Fitzgerald et al., 2013). Taken together, these findings suggest lineups should contain fillers on the higher end of the similarity spectrum. However, two qualifications are worthy of note. First, the authors of the meta-analysis warned that the lineups categorized as having 'high' similarity may not have contained the level of suspect–filler similarity that has been cautioned against. Second, the studies included in the meta-analysis were comprised entirely of adult participants. Whether fillers who resemble the suspect should be used in lineups for child witnesses has yet to be determined.

Children as witnesses

Children may experience difficulty identifying a target person from lineups containing highly similar fillers. Children's underdeveloped capacity to recognize a previously viewed face has been documented in decades of face recognition research. A meta-analysis of face recognition studies revealed that, relative to adults, children are far less adept at discriminating old from new faces (Shapiro & Penrod, 1986). The effect of age on recognizing old faces was particularly pronounced ($d = 1.66$). However, face recognition studies generally differ substantially from a typical eyewitness experience. When the recognition task involves identifying only one or two target persons from a lineup, the difference between children and adults has not always been so apparent.

Pozzulo and Lindsay (1998) conducted a meta-analysis contrasting the lineup identification responses of adults and children. Children were less likely than adults to reject lineups that did not contain the target person, an effect that was especially pronounced when lineup members were presented sequentially. However, children and adults were roughly equivalent when the target was present. Thus, in contrast to the age differences observed in studies using a laboratory-style face recognition paradigm (Shapiro & Penrod, 1986), the meta-analysis of lineup identification studies provided little indication of age-related deficits in children's ability to correctly identify a target.

If children are truly as effective as adults at identifying targets from lineups, perhaps age is not an important factor when considering lineup composition. However, there is good reason to question the finding that children and adults have similar identification abilities. At the time of Pozzulo and Lindsay's (1998) meta-analysis, research exploring age differences in children was still in its early stages. The meta-analysis was based on data from 13 studies, most of which had relatively small sample sizes, raising questions about whether there was sufficient statistical power to detect a difference. For this reason, Pozzulo and Lindsay were tentative in many of their conclusions.

Since Pozzulo and Lindsay's (1998) meta-analysis, the finding that children are particularly prone to making false identifications from target-absent lineups has been replicated in several studies (Havard, Memon, Laybourn, & Cunningham, 2012; Keast, Brewer, & Wells, 2007; Pozzulo & Balfour, 2006; Pozzulo & Dempsey, 2006). By contrast, associations between age and correct identifications have been mixed. In some studies, children have identified their target as accurately as adults (Havard et al., 2012; Pozzulo & Warren, 2003); however, in other studies, children have exhibited an underdeveloped ability to make correct identifications. For example, in one study with a large sample (Keast et al., 2007; n = 616 for target-present lineups), the correct identification rate for adults (40%) was twice the rate for 10- to 13-year-olds (20%). Thus, children do not always perform as well as adults on target-present lineups (also Leippe, Romanczyk, & Manion, 1991). Even if the early meta-analytic finding that children and adults have comparable target-present performance were to be taken at face value, the absence of any factorial manipulations of age and target–filler similarity would preclude a determination that age should not factor into considerations of lineup composition.

At present, very little is known about how lineup member similarity affects children's identification decisions. A review of the literature revealed no direct manipulations of similarity that involved child witnesses. Although authors of child witness research frequently indicated either that fillers resembled the target (Gross & Hayne, 1996; Leippe et al., 1991; Lindsay, Pozzulo, Craig, Lee, & Corber, 1997; Pozzulo & Lindsay, 1997; Pozzulo & Warren, 2003) or that fillers matched the target's description (Brewer & Day, 2005; Brewer, Keast, & Sauer, 2010; Hafstad, Memon, & Logie, 2004; Lowenstein, Blank, & Sauer, 2010; Pozzulo & Balfour, 2006), they made no mention of similarity ratings. In some studies (Memon & Rose, 2002; Pozzulo & Dempsey, 2006; Pozzulo, Dempsey, Crescini, & Lemieux, 2009; Pozzulo, Dempsey, & Wells, 2010), researchers mentioned collecting similarity ratings but did not report the actual ratings. In other studies, researchers collected ratings of distinctiveness rather than similarity (Clifford, Havard, Memon, & Gabbert, 2012; Havard, Memon, Clifford, & Gabbert, 2010; Havard et al., 2012). We only located one study in which quantitative data indicating the similarity between the target and fillers were reported (Keast et al., 2007). Although obviously a within-study similarity manipulation with children would be most useful, the common practice of not reporting similarity ratings makes it difficult to even make comparisons between studies that might have differed in target–filler similarity.

In the experiments that follow, we manipulated similarity in lineup tasks with children to fill this gap in the literature. A critical concern is the impact of highly similar fillers on children's identification of the target. If children have an underdeveloped capacity to discriminate old from new faces, as was suggested in the meta-analysis of face recognition studies (Shapiro & Penrod, 1986), increasing the similarity of lineup members could present an exceedingly difficult task for child witnesses. On the contrary, if children have a fully developed capacity to make correct identifications from target-present lineups, as was suggested in the meta-analysis of lineup identification studies (Pozzulo & Lindsay, 1998), increasing similarity could enhance protection of innocent suspects without substantially reducing correct identifications.

EXPERIMENT 1

In our first experiment, children engaged in a live interaction with a target person. Following a delay, the children were asked to try to identify the target from lineups that contained a suspect who was either guilty (the target) or innocent (an innocent suspect). Fillers of lower or higher resemblance to the target accompanied the suspects. Relative to fillers of lower resemblance to the target, we predicted fillers of higher resemblance would reduce suspect identification rates. More specifically, we predicted higher-similarity fillers would reduce both target selections from target-present lineups and innocent suspect selections from target-absent lineups (Hypothesis 1). The extent to which fillers resemble the target generally does not affect lineup rejections in adults (Fitzgerald et al., 2013), so we hypothesized that similarity would affect only suspect and filler choices (Hypothesis 2).

Method

Participants. We recruited 189 children (M_{age} = 10.06, *SD* = 2.32) from summer science camps. The camps were organized into two groups that differed in age. The younger group comprised 71 children (45 boys) between the ages of 6 and 9 years (*M* = 7.39, *SD* = 1.01). The older group comprised 118 children (75 boys) between the ages of 10 and 14 years (*M* = 11.65, *SD* = 1.07). Five children were excluded from all analyses because they either made multiple selections or indicated they were 'not sure' on the identification task.

Lineups

Lineup members were arranged in a 2 × 3 array. Each lineup member's image was 2.5 in. × 2.1 in. To reduce the likelihood that the target's picture would stand out, the images were printed in grayscale. The suspect's spatial location was counterbalanced across all six lineup positions. A box with the words 'not here' was included at the bottom of the page.

Lineup construction began by recruiting 21 adult judges to make pairwise similarity ratings between the target and 277 potential fillers. All potential fillers were the same race and sex as the target. Judges completed the ratings on a computer in a laboratory setting. Across 277 trials, participants viewed the faces of the target and a potential filler. On each trial, judges received the following instruction: 'Use the scale below to rate the similarity (in terms of physical appearance) between the first person and the second person.' Ratings were made on an 11-point scale (0 = not at all similar; 10 = highly similar). The computer program randomly selected the trial order.

Mean similarity ratings for the set of 277 potential fillers ranged between 0.45 and 5.70 (*M* = 2.27, *SD* = 0.97). We used these ratings to create eight simultaneous lineups: four target-present and four target-absent lineups. The target-present lineups included the target and five fillers. Two of the target-present lineups comprised lower-similarity fillers (similarity ratings: *M* = 1.78 and 1.93; *SD* = 0.35 and 0.11), and two comprised higher-similarity fillers (*M* = 4.44 and 4.76; *SD* = 0.76 and 0.61). Two versions of the lower-similarity and higher-similarity lineups were created for stimulus sampling purposes (Wells & Windschitl, 1999).

We constructed target-absent lineups by taking the target-present lineups and replacing the target with a similar-looking innocent suspect (similarity rating: *M* = 5.65). Thus, target-present and target-absent lineups both contained fillers who were matched to the target’s appearance. Note that the culprit’s appearance would not be known in a police investigation, so our procedure for constructing target-absent lineups would not correspond with a match-to-suspect procedure for an innocent suspect in an actual case. In previous research, using the same fillers in target-present and target-absent lineups led to an underestimation of the innocent suspect misidentification rate (Clark & Tunnicliff, 2001). However, we decided against using different fillers in target-present and target-absent lineups because changing only one aspect of the lineup (i.e., the suspect) provides greater experimental control.

Procedure. For the target event, a 27-year-old man led the children through a 15-minute play session as a special activity during their summer science camp. To encourage the children to attend to the target’s face, he began by introducing himself and outlining the activities he would facilitate. Play session activities for the groups of younger children (6- to

9-year-olds) consisted of mental and physical exercises. The mental exercises were not appropriate for older children, so the 10- to 14-year-olds only completed physical exercises. To keep exposure duration equivalent, the older children did more physical exercises than the younger children.

One day later, research assistants who were blind to the identity of the suspect interviewed the children about the target event. The interviews consisted of two phases: recall and identification. Interviewers were provided with basic instructions on how to interact with children (e.g., dress casually, build rapport, and obtain assent). The interview protocol for the recall phase was unstructured. Rather than providing the interviewers with specific questions to ask, we only instructed the interviewers to obtain the children's full account of the event. After the children had exhausted their memory of the target event, interviewers proceeded to a structured lineup identification task. Before the lineup was presented, interviewers were instructed to make it clear to the children that the visitor may or may not be in the lineup. After a lineup decision was made, the interviewers asked the children for a confidence judgement.[1] After a confidence judgement was obtained, the interviewers thanked the children and gave them a special pen as compensation.

Results

In preliminary analyses, three variables (age group, similarity, and lineup choice) were entered into two hierarchical log-linear (HILOG) analyses: one for target-present lineups and one for target-absent lineups. Both HILOG analyses produced models that were best predicted by a two-factor interaction. Partial association tests showed the two-way interaction between similarity and lineup choice was significant in both HILOG analyses; however, no significant interactions involved the age factor. Accordingly, we collapsed across age groups for all subsequent analyses.

[1] The confidence assessments were beyond the scope of the present research and, thus, are not reported here.

Table 1 presents lineup choice rates in the similarity conditions for target-present and target-absent lineups. Below, we report chi-square tests assessing the association between similarity and lineup choice. To assess similarity effects on each of the three lineup responses, we conducted *z* tests for the difference between two proportions. Cohen's *h*, which is the difference between two arcsine-transformed probabilities, is provided as an effect size metric. The scale used to interpret Cohen's *h* is the same as for Cohen's *d*: small = 0.20, medium = 0.50, and large = 0.80 (Cohen, 1988).

Target present. Similarity was significantly associated with lineup choice in the target-present condition, $\chi^2(2) = 10.53$, $p = .005$. Follow-up analyses indicated similarity affected filler and suspect choices. Specifically, the correct identification rate decreased for higher-similarity relative to lower-similarity lineups, $z = 2.16$, $p = .03$, $h = 0.47$. Conversely, the filler identification rate increased for higher-similarity relative to lower-similarity lineups, $z = 3.04$, $p = .002$, $h = 0.64$. Thus, fillers in the higher-similarity lineups appear to have drawn choices away from the target. Similarity was not associated with lineup rejections, $z = 1.18$, $p = .24$, $h = 0.25$.

Target absent. The association between similarity and lineup choice was also significant in the target-absent condition, $\chi^2(2) = 14.69$, $p < .001$. Follow-up analyses indicated the innocent suspect was more likely to be misidentified from lower-similarity than higher-similarity lineups, $z = 3.48$, $p < .001$, $h = 0.74$, and fillers were more likely to be misidentified from higher-similarity than lower-similarity lineups, $z = 2.65$, $p = .008$, $h = 0.55$. This pattern of results is consistent with that observed with target-present lineups. Thus, regardless of whether the lineup contained the target person or an innocent suspect, higher-similarity fillers drew choices away from the suspect. Also consistent with target-present lineups, similarity was not associated with lineup rejections, $z = 0.50$, $p = .61$, $h = 0.10$.

Diagnosticity. Diagnosticity ratios represent the relative likelihood that a suspect identification is indicative of the suspect's guilt (Wells & Lindsay, 1980). Diagnosticity is calculated by dividing the target choice rate by the innocent suspect choice rate. If the ratio exceeds unity (i.e., 1.00), a suspect identification from that lineup is more likely to indicate guilt than innocence. The diagnosticity ratio for higher-similarity lineups (1.60; 95% CI [0.28, 9.14]) was larger than the diagnosticity ratio for lower-similarity lineups (0.76; 95% CI [0.38, 1.54]); however, the confidence intervals associated with the diagnosticity ratios in both similarity conditions overlapped with 1.00, suggesting suspect identifications were not reliably diagnostic of guilt in any of the lineups.

Discussion

The similarity manipulation effectively influenced children's identification responses. Consistent with Hypothesis 1, higher similarity was associated with a reduction in suspect identifications. For the lower-similarity lineups, the vast majority of children's positive identifications were of the suspect. As similarity increased, however, identifications shifted from the suspect to the fillers. Given that this pattern was uninfluenced by whether the target was present or absent, increasing similarity had a negative effect in the case of target-present lineups (decreased correct identifications) and a positive effect in the case of target-absent lineups (decreased innocent suspect misidentifications). These effects were uninfluenced by the children's age, suggesting a robust effect of lineup member similarity.

The data also support Hypothesis 2, which stated that similarity would only affect suspect and filler choices. Lineup rejection rates, which were notably higher than typically observed for child witnesses, were unaffected by lineup similarity. Given that rejection rates were also unaffected by the target's presence, children's reluctance to choose seemed to be indicative of a global shift toward conservative responding. In other words, given that children rejected the lineup at high rates regardless of whether it was a correct or incorrect decision, we are not convinced that children were always rejecting the target-absent lineups because they knew the target was not present. Rather, we suspect that some aspect of the experimental procedure led children to develop an inclination to say ‘not here’ rather than to choose a lineup-member. For example, children have previously shown an increased tendency to reject lineups when information-processing conditions were poor (Fitzgerald, Price, & Connolly, 2012). Although the reason for children’s conservatism in Experiment 1 is not entirely clear, it is possible that the structure of the target event did not sufficiently motivate children to attend to the target’s face. This potential limitation was addressed in Experiment 2.

EXPERIMENT 2

In Experiment 1, the lineups containing higher-similarity fillers posed a particularly difficult task for the children. However, the correct identification rate in the lower-similarity condition was also not as high as in some previous investigations with children (Pozzulo & Lindsay, 1998). In addition to the composition of the lineups, some procedural artifacts in Experiment 1 might have contributed to the low accuracy rates. For example, children completed a recall task prior to completing the identification task. During the recall task, some children described the target person’s appearance. Therefore, it is possible that this verbal description overshadowed their visual memory of the target person (Schooler & Engstler-Schooler, 1990). The nature of the target event in Experiment 1 also might have influenced the strength of the children’s memory. For example, children encountered the target during a live interaction, and although the children were encouraged to focus on his face during the introductory phase of the encounter, their focus may have been diverted to the activities for the remainder of the event.

In our second experiment, we changed three aspects of the procedure to increase the children’s accuracy rates to a level more typically observed in eyewitness identification research. First, children completed the identification task prior to any recall tasks, thus eliminating any possibility of verbal overshadowing. Second, participants viewed the target person on video rather than in a live interaction. Throughout the duration of the video, a close-up of the target's shoulders and head was displayed. Thus, in addition to controlling the length of target exposure and the content of the target event, the video encouraged participants to focus on the target person's face. Third, participants were warned of an upcoming memory task. Although participants were not specifically warned that they would be making an identification, they were informed that they would be required to remember the words the target person was saying. These procedures allowed us to examine the effect of higher-similarity fillers for an event that was encoded under favorable information-processing conditions.

In addition to these procedural changes, we included an adult group for comparison. Although the effects observed in Experiment 1 corresponded with the pattern of results previously observed with adults, the difference in accuracy between children in the present study and adults in previous studies is noteworthy. For example, in the meta-analysis examining similarity effects in adults, more than 40% identified the target from a high-similarity lineup (Fitzgerald et al., 2013). This rate contrasts starkly with the 7% correct identification rate associated with children in the higher-similarity condition of Experiment 1. Although this disparity could be interpreted as an indication that children are less proficient than adults on lineups containing higher-similarity fillers, age was not the only factor that differed between Experiment 1 and previous research with adults. Thus, differences in experimental procedures and materials could account for the discrepant effects of higher-similarity lineups on children and adults.

In Experiment 2, we examined the effect of similarity on identification responses of children and adults using comparable procedures and materials. We predicted that the pattern of effects observed with children in Experiment 1 would be replicated in Experiment 2. In particular, we predicted that increased similarity would decrease children's suspect identification rates and increase their filler identification rates (Hypothesis 1). Given children's poor performance on the higher-similarity lineups in Experiment 1 and previous research showing adults can identify targets from high-similarity lineups, we hypothesized that correctly identifying the target person from higher-similarity lineups would be more difficult for children than for adults (Hypothesis 2).

Method

Participants. The sample consisted of 164 children and 172 adults. Children (110 boys) were recruited from a summer science camp. The children's ages ranged from 6 to 12 years (*M* = 8.92, *SD* = 1.41). Adults (39 men) were recruited from undergraduate psychology courses. The adults' ages ranged from 17 to 56 years (*M* = 20.40, *SD* = 3.78). For compensation, children received a special pen, and adults received partial course credit. In addition to differing in age, the child and adult groups differed in their male-to-female ratios. In particular, the child group consisted mostly of boys (67%), and the adult group consisted mostly of women (77%), $\chi^2(1) = 67.05$, $p < .001$. However, male (65%) and female (68%) children were comparable in identification accuracy, $\chi^2(1) = 0.07$, $p = .79$, and male (67%) and female (65%) adults were also comparable in identification accuracy, $\chi^2(1) = 0.05$, $p = .82$. Therefore, although male and female participants were not represented equally in the two age groups, sex was not associated with identification accuracy.

Lineups

Eight lineups were constructed to produce the same lineup manipulations (target presence, similarity, and version) employed in Experiment 1. All lineups contained six members who were presented simultaneously in a 2 × 3 array. A 'not here' box was positioned at the bottom of the page. The lineup member images were 2.5 in. × 2.1 in. in size and printed in grayscale. The position of the suspect was counterbalanced across all six positions.

Adult judges (*n* = 35), who were independent from both of the main experiments, made pairwise similarity ratings between the target person and 200 potential fillers of the same race and sex. Judges, who provided these ratings through an online survey, were given the following instructions: ‘In terms of physical appearance, how similar are these two individuals?’ The online survey program randomly selected the order of the trials. The similarity judgements were made on a scale from 1 (highly dissimilar) to 10 (highly similar). Similarity ratings for the set ranged from 1.49 to 6.06 (*M* = 3.51, *SD* = 0.78). Filler ratings in the lower-similarity lineups (Version 1: *M* = 2.38; *SD* = 0.36; Version 2: *M* = 2.53; *SD* = 0.17) were lower than in the higher-similarity lineups (Version 1: *M* =4.73; *SD* = 0.24; Version2: *M* = 4.89; *SD* = 0.51). In target-absent lineups, the target was replaced with a similar-looking innocent suspect (*M* = 6.06).

General procedure. For all participants, the experimental procedure consisted of a magic show performed by a woman and a series of video clips depicting a man. The man (29 years old) in the video served as the target person for the identification task. Participants did not complete an identification task for the woman. In total, five clips of the target person were shown. All clips depicted a close-up of the target person (i.e., only his shoulders and head were in view), and the target person always looked directly at the camera. In the introductory clip (28 seconds), the target person informed viewers that he would be reading words to them that will need to be remembered and that a magician will be performing a trick after each word list. In each of the four word-list clips (23 seconds per clip), the target person read 14 words that were all semantically related to a word that was absent from the list. These lists were used for an unrelated study examining the Deese–Roediger–McDermott paradigm in children (Deese, 1959; Roediger & McDermott, 1995). The target person wore four different types of hats while reading the word lists, which were used to cue participants to specific lists. The target person did not wear a hat for the introductory clip.

All participants viewed the video clips in groups and completed the identification task the following day. The lineups were manipulated in terms of target presence and target– filler similarity. Although the lineup administrators were not blind to the target person's identity, we took steps to ensure they did not know whether the target was present or absent from the lineup. Specifically, the administrators received the lineups in opaque folders that remained closed until immediately before the identification task, at which point they opened the folder so that it was faced toward the participant and outside of the administrator's view. This was designed to prevent any unintentional influences of the administrator on the participant's identification decision. Prior to administering the lineup, participants were informed that the man from the videos may be there or may not be there. They were further instructed that they did not need to choose any of the lineup members. Participants indicated their identification choice either by stating the number associated with a lineup member they believed to be the target or by stating 'not here' if they believed the target was absent. No feedback was provided after the identification. The participants subsequently indicated how sure they were about their identification decision.[2]

Procedure for children. A female magician visited the children's summer science camp and performed a live magic show. The science camps were held at a university campus, and the children watched the magic show in groups in a university classroom. The video clips were projected onto a large screen, ensuring all children had a good view of the target person. After the introductory clip, the magician discussed some safety instructions and instructed children to prepare for the first word list. After children viewed the first word-list clip, the

[2] Confidence assessments are not related to the present hypotheses and, thus, are not reported here.

magician performed her first trick. This sequence was repeated for the remaining word-list clips and magic tricks.

On the following day, interviewers met with children individually to administer the lineups. After obtaining assent, interviewers reminded children about the man from the video. Interviewers subsequently informed children that they would be viewing some pictures and would be asked if the man from the video was in any of the pictures, followed by a warning that the man's picture may or may not be present. Upon presentation of the lineup, children were instructed to report the number associated with their choice (or say 'not here'). After children made an identification decision, the interviewers closed the folder and recorded the decision on a separate response sheet. The interviewers then proceeded to ask the children how sure they were about the identification decision, followed by recall and recognition questions about the magic show and word lists.[3] The interviewers subsequently thanked the children and awarded them a special pen.

Procedure for adults. Adults viewed a video-recorded version of the magic show. A researcher visited multiple undergraduate classes and projected a video that alternated between the five clips of target person and five clips of the magic show (safety instructions and four tricks) onto a large screen, thus simulating the children's experience. Prior to the video's presentation, the researcher informed students that they would be watching a magic show. After the video's completion, the researcher explained that children had viewed the same magic show and their memory for the word lists had been tested. The researcher explained that we were recruiting adults for a comparison group. The researcher then provided the location and times at which interested students could participate for partial

[3] Children's memory of the magic show and word lists is not related to the present hypotheses and, thus, is not reported here.

course credit. The researcher emphasized that students could only participate on the following day.

The identification task was conducted in a laboratory the day after the video presentation. Upon arrival, students were informed that we were actually interested in whether they could identify the person who read the word lists, rather than their memory of the words. After providing informed consent, students were directed to an identification testing room. The lineups and the lineup procedure were identical to those used with children. After completion of the identification and confidence procedures, the administrator thanked the adult for participating. Adults were not asked about the content of the word lists or the magic show.

Results

Target present. A 2 (child vs. adult) × 2 (lower similarity vs. higher similarity) × 3 (target identification vs. filler identification vs. lineup rejection) HILOG analysis revealed a significant three-way interaction, $\chi^2(2) = 6.70$, $p = .04$, suggesting the most predictive model incorporated all three factors (Table 2). A two-way interaction was also significant, $\chi2(5) = 16.68$, $p = .005$. Partial association tests indicated a significant interaction between similarity and lineup choice, $\chi^2(2) = 11.58$, $p = .003$, and a marginal interaction between age group and lineup choice, $\chi^2(2) = 4.67$, $p = .10$.

Tests of simple effects indicated that similarity's effect on correct identifications depended on age group. Children identified the target at a significantly higher rate from lower-similarity lineups (74%) than from higher-similarity lineups (48%), $z = 2.49$, $p = .01$, $h = 0.54$. By contrast, adults identified the target from lower-similarity (76%) and higher-similarity (74%) lineups at comparable rates, $z = 0.12$, $p = .90$, $h = 0.03$. For higher-similarity lineups, the correct identification rate was significantly higher for adults than for children, $z = 2.56$, $p = .01$, $h = 0.55$. For lower-similarity lineups, the correct identification rates for children and adults were comparable, $z = 0.20$, $p = .83$, $h = 0.05$. Thus, children only performed as well as adults on the target-present lineups that contained lower-similarity fillers.

For both children and adults, witnesses identified fillers more frequently in the higher-similarity condition than in the lower-similarity condition. The effect size for children (25% vs. 0%, $h = 1.05$) was much larger than the effect size for adults (15% vs. 7%, $h = 0.25$); however, low frequency counts in the lower-similarity conditions precluded significance tests examining associations between similarity and filler identification rates. In the higher-similarity condition, the filler identification rate was 10% higher for children than for adults; however, the difference was not significant, $z = 1.08$, $p = .28$, $h = 0.24$. Overall, filler identification rates for adults (11%) and children (13%) were comparable.

We found a significant association between age and incorrect rejection rates. In particular, the rejection rate was significantly higher for children (27%) than for adults (14%), $z = 2.04$, $p = .04$, $h = 0.32$. Although the difference in rejection rates between children and adults was larger for higher-similarity lineups (27% vs. 10%, respectively) than for lower-similarity lineups (26% vs. 17%, respectively), both similarity conditions yielded the same age-related trend. Consistent with Experiment 1, similarity was not associated with lineup rejections.

Target absent. A 2 (child vs. adult) × 2 (lower similarity vs. higher similarity) × 3 (innocent suspect identification vs. filler identification vs. lineup rejection) HILOG analysis revealed a significant two-way interaction, $\chi^2(5) = 16.55$, $p = .005$ (Table 2). The three-way interaction was not significant, $\chi^2(2) = 0.73$, $p = .70$. Partial association tests revealed a significant interaction between age group and lineup choice, $\chi^2(2) = 9.40$, $p = .009$, and a significant interaction between similarity and lineup choice, $\chi^2(2) = 7.30$, $p = .03$.

The interaction between age group and lineup choice was indicative of differences in filler identifications and lineup rejections. Although children and adults identified the innocent suspect at comparable rates (24% vs. 22%, respectively), the filler identification rate was higher for adults (22%) than for children (6%), $z = 3.04$, $p = .002$, $h = 0.47$, and the correct rejection rate was marginally higher for children (70%) than for adults (57%), $z = 1.90$, $p = .06$, $h = 0.29$. The influence of age on filler selections and lineup rejections was consistent for lower-similarity and higher-similarity conditions.

The interaction between similarity and lineup choice was primarily driven by differences in innocent suspect misidentifications. Overall, witnesses selected the innocent suspect more frequently from lower-similarity lineups (31%) than from higher-similarity lineups (15%), $z = 2.50$, $p = .01$, $h = 0.38$. Witnesses identified fillers and correctly rejected the lineup more frequently from higher-similarity lineups than from lower-similarity lineups; however, both differences were nonsignificant (filler: 18% vs. 11%, $z = 1.43$, $p = .15$, $h = 0.21$; rejection: 67% vs. 59%, $z = 1.11$, $p = .27$, $h = 0.17$). These response patterns were consistent for children and adults.

Diagnosticity. In the lower-similarity conditions, the extent to which suspect identifications were diagnostic of guilt was similar for adults (2.67; 95%CI [1.64, 4.38]) and children (2.23; 95%CI [1.38, 3.61]). By contrast, in the higher-similarity conditions, adults' suspect identifications (4.88; 95%CI [2.41, 9.91]) were considerably more diagnostic of guilt relative to children's suspect identifications (3.34; 95%CI [1.50, 7.46]).

Discussion

Experiment 2 demonstrated that children are particularly sensitive to differences in lineup member similarity. Accuracy rates in Experiment 2 were generally higher than in Experiment 1, suggesting the video-recorded event was effective in focusing participants' attention on the target and increasing memory strength. In support of Hypothesis 1, higher-similarity fillers led to a reduction in children's correct identification rate. Thus, even with the more favorable encoding conditions and the resulting elevation in memory strength, children experienced difficulty identifying the target from lineups with higher-similarity fillers. By contrast, adults were largely unaffected by the increase in similarity on target-present lineups. The absence of an effect of similarity on adults' target-present performance corresponds with a recent meta-analysis that revealed comparable correct identification rates for adults on moderate-similarity and high-similarity lineups (Fitzgerald et al., 2013).

Hypothesis 2 was supported by a significantly higher correct identification rate for adults than for children on the higher-similarity lineups. Approximately half of the children were unable to identify the target from higher-similarity lineups. These accuracy rates were substantially lower than the rate for adults in the higher-similarity condition, who identified the target three quarters of the time. Thus, the higher-similarity fillers only posed a problem for children's target identifications. Although higher-similarity fillers negatively affected children's performance on target-present lineups, increasing similarity generally had a desirable effect on responses to the target-absent lineups. For adults and children, the innocent suspect was less likely to be identified from higher-similarity lineups than from lower-similarity lineups. Thus, higher-similarity fillers led to a benefit in children and adults on target-absent lineups and led to a cost in children but not adults on target-present lineups.

Although the results were generally consistent with our stated hypotheses, we were surprised by the age differences in choosing behavior. We made no hypotheses about whether children or adults would be more likely to reject the lineup, but had we done so, our predictions almost certainly would have been wrong. Children are notorious for their inclination to choose (e.g., Pozzulo & Lindsay, 1998); however, relative to adults in Experiment 2, children rejected the lineups at higher rates. This was true regardless of whether the target was present or absent, suggesting some aspect of our procedure was leading children to adopt a relatively strict decision criterion for making a positive identification.

Children in Experiment 1 also frequently rejected the lineups, which we interpreted as an indication of low memory strength caused by poor encoding conditions. This interpretation was based on the finding that children in Experiment 1 made a high number of both correct (64%) and incorrect (55%) rejections. By contrast, children in Experiment 2 were far more likely to correctly reject the lineup (70%) than they were to incorrectly reject the lineup (27%). Thus, children's rejections in Experiment 2 seem indicative of memory strength rather than a global shift in decision threshold that favors lineup rejections.

A number of factors may have contributed to children's relatively strong target-absent performance in Experiment 2. We included several procedures that have previously been shown to improve children's target-absent performance such as warning children the target may or may not be present (Keast et al., 2007), instructing lineup administrators to dress casually (Lowenstein et al., 2010), and including a salient rejection option in the lineup (Zajac & Karageorge, 2009). However, such procedures are commonly used in studies comparing child and adult witnesses, and to our knowledge, this is the first report of children rejecting the lineup at a higher rate than adults.

The key difference between the present and previous research was our focus on similarity. It is possible that some children were overwhelmed with the homogeneity of the lineup members and opted to reject the lineup rather than make a difficult choice. Consistent with this explanation, the difference in rejection rates between children and adults was greater in the higher-similarity conditions than in the lower-similarity conditions. However, these differences were relatively small in magnitude. Given that our only experimentally manipulated variable in the target-absent lineups (similarity) was not significantly associated with age differences in lineup rejections, we cannot make any causal attributions about children's surprisingly strong performance. Therefore, we recommend conducting future research containing experimental manipulations of different aspects of our procedure to investigate this issue further.

GENERAL DISCUSSION

Wells and colleagues (Luus & Wells, 1991; Wells et al., 1993) cautioned against using fillers who resemble the suspect too strongly out of concern that highly similar lineup members would impede culprit identifications. Given the results of the present research, this concern seems particularly relevant for children. In Experiment 1, children had an exceptionally low rate of correct identifications from higher-similarity lineups. In Experiment 2, children showed some capacity to make correct identifications from higher-similarity lineups, but the encoding conditions were likely much more favorable than would be the case in an actual eyewitness encounter. Even under these ideal information-processing conditions, children were significantly less likely to make a correct identification on higher-similarity relative to lower-similarity lineups. By contrast, increasing lineup member similarity had no effect on adults' correct identification rates. These findings clearly demonstrate that the type of lineups that would be appropriate for adults poses a task too difficult for children.

Although our findings show that children have an underdeveloped capacity to make identifications from lineups containing fillers on the higher end of the similarity spectrum, further research is required to determine just how similar is too similar. We used the labels 'lower similarity' and 'higher similarity' to reflect the fact that our manipulations led to relative differences in target–filler similarity. In the absence of an objective metric for establishing the similarity between two persons, we are not in a position to comment on the degree of similarity in the higher-similarity lineups on any kind of absolute scale. However, the ratings we obtained were comparable with those obtained in previous research (e.g., Sauer, Brewer, & Weber, 2008).

The extent to which the lineups we created correspond with those used by police intending to create a ‘fair’ lineup is also worthy of consideration. When police detectives have been instructed to select fillers who resemble the suspect (Juslin et al., 1996; Tunnicliff & Clark, 2000, Experiment 1), the suspect–filler similarity ratings for their lineups were lower than the ratings for the higher-similarity lineups in the present research. Tunnicliff and Clark, who did not provide the detectives in their study with instructions about precisely how similar the fillers should be, suggested the detectives likely did not select the most similar fillers available. Thus, police may use lineups with less similarity than those employed in the present research.

Further evidence that police tend to use lower-similarity lineups comes from Valentine and Heaton (1999), who analyzed lineups from actual cases and found that the fillers often did not match the witness description of the culprit as well as did the suspect. It is important to note, however, that in the absence of strict guidelines governing lineup construction procedures, the composition of lineups employed across police agencies will be subject to substantial variation. In the United States, for example, police responding to a survey reported three methods of filler selection: 19% match fillers to the witness description of the culprit, 31% match fillers as closely as possible to suspect’s appearance, and 50% match fillers to a general description of the suspect (Police Executive Research Forum, 2013). Our research suggests at least one of these methods would not be ideal for use with child witnesses.

Final considerations

Our findings suggest the age of the witness should be taken into account when constructing lineups. If the witness is a child, police should not fill the lineup with members as similar as those used to fill a lineup for adults. However, police investigators need to interpret this recommendation carefully. For investigators who routinely select the most similar fillers available, an appropriate interpretation of our findings would be to continue their current practice for adults and to change their current practice for children by selecting less similar fillers. Given that only 31% of police investigators reported selecting the most similar fillers available (Police Executive Research Forum, 2013), most investigators would likely benefit from a different course of action. For the 69% of investigators who do not routinely select the most similar fillers available, an appropriate interpretation of our findings would be to change their current practice for adults by selecting more similar fillers and to not make any changes to their current practice with children. This is a critical distinction because if the latter group of investigators were to lower their standard of lineup member similarity for children, the resulting lineup could contain a high degree of bias toward an innocent suspect.

References

Brewer, N., & Day, K. (2005). The confidence-accuracy and decision latency-accuracy relationship in children's eyewitness identification. *Psychiatry, Psychology and Law, 12,* 119-128. doi: 10.1375/pplt.2005.12.1.119

Brewer, N., Keast, A., & Sauer, J. D. (2010). Children's eyewitness identification performance: Effects of a not sure response option and accuracy motivation. *Legal and Criminological Psychology, 15,* 251-277. doi: 10.1348/135532509X474822

Clark, S. E., & Tunnicliff, J. L. (2001). Selecting lineup foils in eyewitness identification experiments: Experimental control and real-world simulation. *Law and Human Behavior, 25,* 199-216. doi:10.1023/A:1010753809988

Clifford, B. R., Havard, K., Memon, A., & Gabbert, F. (2012). Delay and age effects on identification accuracy and confidence: An investigation using a video identification parade. *Applied Cognitive Psychology, 26,* 130-139. doi: 10.1002/acp.1804

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Darling, S., Valentine, T., & Memon, A. (2008). Selection of lineup foils in operational contexts. *Applied Cognitive Psychology, 22,* 159-169. doi: 10.1002/acp.1366

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58,* 17-22. doi: 10.1037/h0046671

Fitzgerald, R. J., Price, H. L., & Connolly, D. A. (2012). Anxious and nonanxious children's face identification. *Applied Cognitive Psychology, 26,* 585-593. doi: 10.1002/acp.2833

Fitzgerald, R. J., Price, H. L., Oriet, C., & Charman, S. D. (2013). The effect of suspect-filler similarity on eyewitness identification decisions: A meta-analysis. *Psychology, Public Policy, and Law, 19,* 151-164. doi: 10.1037/a0030618

Gross, J., & Hayne, H. (1996). Eyewitness identification by 5- to 6-year-old children. *Law and Human Behavior, 20,* 359-373. doi: 10.1007/BF01499028

Hafstad, G. A., Memon, A., & Logie, R. (2004). Post-identification feedback, confidence and recollections of witnessing conditions in child witnesses. *Applied Cognitive Psychology, 18,* 901-912. doi: 10.1002/acp.1037

Havard, K., Memon, A., Clifford, B. R., & Gabbert, F. (2010). A comparison of video and static photo lineups with child and adolescent witnesses. *Applied Cognitive Psychology, 24,* 1209-1221. doi: 10.1002/acp.1645

Havard, K., Memon, A., Laybourn, P., & Cunningham, C. (2012). Own-age bias in video lineups: a comparison between children and adults. *Psychology, Crime & Law, 18,* 929-944. doi:10.1080/1068316X.2011.598156

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence-accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1304-1316. doi: 10.1037/0278-7393.22.5.1304

Keast, A., Brewer, N., & Wells, G. L. (2007). Children’s metacognitive judgements in an eyewitness identification task. *Journal of Experimental Child Psychology, 97,* 286-314. doi: 10.1016/j.jecp.2007.01.007

Leippe, M., Romanczyk, A., & Manion, A. P. (1991). Eyewitness memory for a touching experience: Accuracy differences between child and adult witnesses. *The Journal of Applied Psychology, 76,* 367–379.

Lindsay, R. C. L., Martin, R., & Webber, L. (1994). Default values in eyewitness descriptions: A problem for the match-to-description lineup foil selection strategy. *Law and Human Behavior*, *18,* 527–541. doi: 10.1007/BF01499172

Lindsay, R. C. L., Pozzulo, J. D., Craig, W., Lee, K., & Corber, S. (1997). Simultaneous lineups, sequential lineups, and showups: Eyewitness identification decisions of adults and children. *Law and Human Behavior, 21,* 391-404. doi: 10.1023/A:1024807202926

Lindsay, R. C. L., & Wells, G. L. (1980). What price justice? Exploring the relationship of lineup fairness to identification accuracy. *Law and Human Behavior, 4,* 303-313. doi: 10.1007/BF01040622

Lowenstein, J. A., Blank, H., & Sauer, J. D. (2010). Uniforms affect the accuracy of children's eyewitness identification decisions. *Journal of Investigative Psychology and Offender Profiling, 7,* 59–73. doi: 10.1002/jip.104

Luus, C. A. E., & Wells, G. L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior, 15,* 43-57. doi: 10.1007/BF01044829

Memon, A., & Rose, R. (2002). Identification abilities of children: Does a verbal description hurt face recognition? *Psychology, Crime & Law*, *8,* 229-242. doi: 10.1080/10683160210000015772

Police Executive Research Forum. (2013). *A National Survey of Eyewitness Identification Processes in Law Enforcement Agencies*. Washington, DC: Author. Retrieved from http://policeforum.org/library/eyewitness-dentification/NIJEyewitnessReport.pdf

Pozzulo, J. D., & Balfour, J. (2006). Children's and adults' eyewitness identification accuracy when a culprit changes his appearance: Comparing simultaneous and elimination lineup procedures. *Legal and Criminological Psychology, 11,* 25-34. doi: 10.1348/135532505X52626

Pozzulo, J. D., & Dempsey, J. (2006). Biased lineup instructions: Examining the effect of pressure on children's and adults' eyewitness identification accuracy. *Journal of Applied Social Psychology, 36,* 1381-1394. doi: 10.1111/j.0021-9029.2006.00064.x

Pozzulo, J. D., Dempsey, J. L., Crescini, C. & Lemieux, J. M. T. (2009). Examining the relation between eyewitness recall and recognition for children and adults. *Psychology, Crime & Law, 15,* 409-424. doi: 10.1080/10683160802279625

Pozzulo, J. D., Dempsey, J. L., & Wells, K. (2010). Does lineup size matter with child witnesses. *Journal of Police and Criminal Psychology, 25,* 22–26. doi: 10.1007/s11896-009-9055-x

Pozzulo, J. D., & Lindsay, R. C. L. (1997). Increasing correct identifications by children. *Expert Evidence, 5,* 126-132. doi: 10.1023/A:1008875802767

Pozzulo, J. D., & Lindsay, R. C. L. (1998). Identification accuracy of children versus adults: A meta-analysis. *Law and Human Behavior, 22,* 549-570. doi: 10.1023/A:1025739514042

Pozzulo, J. D., & Lindsay, R. C. L. (1999). Elimination lineups: An improved identification procedure for child eyewitnesses. *Journal of Applied Psychology, 84,* 167-176. doi: 10.1037/0021-9010.84.2.167

Pozzulo, J. D., & Warren, K. L. (2003). Descriptions and identifications of strangers by youth and adult eyewitnesses. *Journal of Applied Psychology, 88,* 315-323. doi: 10.1037/0021-9010.88.2.315

Roediger, H. L. III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning Memory and Cognition, 21,* 803-814. doi: 10.1037/0278-7393.21.4.803

Sauer, J. D., Brewer, N., Weber, N. (2008). Multiple confidence estimates as indices of eyewitness memory. *Journal of Experimental Psychology: General, 137,* 528-547.

doi: 10.1037/a0012712

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology, 22,* 36-71. doi: 10.1016/0010-0285(90)90003-M

Shapiro, P., & Penrod, S.D. (1986). A meta-analysis of the facial identification literature. *Psychological Bulletin, 100,* 139-156.

Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement*. Washington, DC: National Institute of Justice.

Tunnicliff, J. L., & Clark, S. E. (2000). Selecting foils for identification lineups: Matching suspects or descriptions? *Law and Human Behavior, 24,* 231-258. doi: 10.1023/A:1005463020252

Valentine, T., & Heaton, P. (1999). An evaluation of the fairness of police line-ups and video identifications. *Applied Cognitive Psychology, 13,* S59-S72. doi: 10.1002/(SICI)1099-0720(199911)13:1+<S59::AID-ACP679>3.0.CO;2-Y

Wells, G. L., & Lindsay, R. C. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin, 3,* 776-784. doi: 10.1037/0033-2909.88.3.776

Wells, G. L., Memon, A., & Penrod, S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest, 7,* 45-75. doi: 10.1111/j.1529-1006.2006.00027.x

Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distractors for eyewitness lineups. *Journal of Applied Psychology, 78,* 835-844. doi: 10.1037/0021-9010.78.5.835

Wells, G. L., & Turtle, J. W. (1986). Eyewitness identification: The importance of lineup models. *Psychological Bulletin, 99,* 320-329. doi: 10.1037/0033-2909.99.3.320

Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin, 25,* 1115-1125. doi: 10.1177/01461672992512005

Wogalter, M. S., Malpass, R. S., & McQuiston, D. E. (2004). A national survey of U. S. police on preparation and conduct of identification lineups. *Psychology, Crime & Law, 10*, 69-82. doi: 10.1080/106831604100016418

Zajac, R., & Karageorge, A. (2009). The wildcard: A simple technique for improving children's target-absent lineup performance. *Applied Cognitive Psychology, 23,* 358 368. doi:10.1002/acp.1511

Table 1

Lineup response probabilities for children (ages 6-14) in Experiment 1

Target	**Similarity**	**Choice**			***n***
		Target	**Filler**	**Reject**	
Present	Lower	.23	.16	.61	44
	Higher	.07	.44	.49	45
	Overall	.15	.30	.55	89
		Innocent	**Filler**	**Reject**	
Absent	Lower	.30	.09	.62	47
	Higher	.04	.30	.67	48
	Overall	.17	.19	.64	95

Table 2

Identification response probabilities for children (ages 6-12) and adults in Experiment 2

Target	**Age Group**	**Similarity**	**Choice**			***n***
			Target	**Filler**	**Reject**	
Present	Children	Lower	.74	.00	.26	39
		Higher	.48	.25	.27	44
		Overall	.60	.13	.27	83
	Adults	Lower	.76	.07	.17	41
		Higher	.74	.15	.10	39
		Overall	.75	.11	.14	80
			Innocent	**Filler**	**Reject**	
Absent	Children	Lower	.33	.03	.64	39
		Higher	.14	.10	.76	42
		Overall	.24	.06	.70	81
	Adults	Lower	.28	.17	.54	46
		Higher	.15	.26	.59	46
		Overall	.22	.22	.57	92