Law and Human Behavior, Vol. 28, No. 3, June 2004 (© 2004)

Lineup Identification by Children: Effects of Clothing Bias

Alejo Freire,[1,5] **Kang Lee,**[2] **Karen S. Williamson,**[3] **Sarah J. E. Stuart,**[4] **and R. C. L. Lindsay**[4]

This study examined effects of clothing cues on children's identification accuracy from lineups. Four- to 14-year-olds (n = *228) saw 12 video clips of individuals, each wearing a distinctly colored shirt. After watching each clip children were presented with a target-present or target-absent photo lineup. Three clothing conditions were included. In 2 conditions all lineup members wore the same colored shirt; in the third,* biased *condition, the shirt color of only one individual matched that seen in the preceding clip (the* target *in target-present trials and the* replacement *in target-absent trials). Correct identifications of the target in target-present trials were most frequent in the biased condition, whereas in target-absent trials the biased condition led to more false identifications of the target replacement. Older children were more accurate than younger children, both in choosing the target from target-present lineups and rejecting target-absent lineups. These findings suggest that a simple clothing cue such as shirt color can have a significant impact on children's lineup identification accuracy.*

KEY WORDS: eyewitness testimony; face recognition; development, & clothing bias.

When individuals witness a crime police may present them with photographic lineups to identify the perpetrators. This task is complicated by the fact that police may arrest suspects who are innocent. In other words, a lineup may include a perpetrator, known as a target-present lineup, or may not, known as a target-absent lineup. Lineup procedures should thus be designed to maximize the ability of witnesses to correctly identify perpetrators, minimize the likelihood of picking innocent suspects, and reduce potential sources of bias (Lindsay & Wells, 1985; Wells et al., 2000; Wells, Rydell, & Seelau, 1993).

[1]McMaster University, Hamilton, Ontario, Canada.
[2]Department of Psychology, University of California at San Diego, La Jolla California.
[3]University of Guelph, Guelph, Canada.
[4]Queen's University, Canada.
[5]To whom correspondence should be addressed at Department of Psychology, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1; e-mail: freirea@mcmaster.ca.

Bias may affect a witness' response by either subtly or overtly encouraging the choice of one lineup member relative to the others. Numerous factors may introduce bias into a lineup procedure, including whether photos are seen simultaneously or sequentially, the specific instructions provided prior to and during presentation of the lineup, and the similarity of foils relative to the target in terms of physical and facial resemblance (e.g., Lindsay & Wells, 1980; Lindsay & Wells, 1985; Malpass & Devine, 1981). For example, a person may feel pressured to identify the person in a photo as the perpetrator if it is the only photo presented to her by the police (referred to as a *show up*; Dysart, Lindsay, MacDonald, & Wicke, 2002). Similarly, if the witnesses are instructed that the perpetrator is present in the lineup and the task is simply to identify the correct individual, it follows that they will feel pressured to make a choice (Malpass & Devine, 1981; Steblay, 1997).

It has been argued that one source of bias is an increase in the use of relative judgment strategies when making a decision from a lineup (e.g., Lindsay & Wells, 1985; Wells, 1984). That is, rather than making an absolute judgment that a lineup member is the culprit, a relative judgment is made wherein the person most resembling the suspect is identified. Although this strategy should lead to correct identifications when the guilty party is present in the lineup, it leads to the false identification of innocent individuals when the culprit is not included in the lineup. Lineup biases lead to incorrect selections either by encouraging the use of relative judgments or by making the innocent suspect stand out such that witnesses using relative judgments are more likely to select the suspect. As a result, any factor that makes the suspect stand out in a lineup results in a biased procedure.

The clothing worn by lineup members is a factor with the potential to bias a lineup, but it has not been studied extensively (Lindsay, Nosworthy, Martin, & Martynuk, 1994; Lindsay, Wallbridge, & Drennan, 1987). To the extent that the suspect stands out because his or her clothing is a better match than the clothes worn by other lineup members to the clothes worn in commission of the crime, the lineup will be biased. The frequency of clothing bias in real world lineups is unknown but one author (RCLL) has found clothing bias in approximately one third of the criminal cases in which he has consulted on lineup issues. This suggests that the problem is not uncommon.

The limited evidence related to this issue suggests that clothing cues can indeed contribute to misidentifications. Lindsay et al. (1987) had adult participants witness a staged theft and later presented them with lineups. In the target present conditions the lineup comprised pictures of the perpetrator and five foils. In the target-absent conditions the lineup was made up of photos of the same five foils and, in lieu of the target, a *replacement*. In addition to being either a target-present or target-absent lineup, each lineup included one of three clothing manipulations. The first clothing condition was the *usual* condition, where each lineup member was dressed differently and with clothing unlike that worn during the crime. This condition represented usual police practice. The second was the *dressed alike* condition, where all lineup members were similarly dressed in either a white lab coat or common sweatshirt. Dressing all lineup members alike assured no clothing bias. The condition in which all lineup members wore clothing similar to that worn during the crime (the sweatshirt) may have provided a cue associated only with the criminal and thus increased correct

identification rates. The lab coat condition represented an easily accomplished control of clothing cues that police could adopt. The last condition was the *biased* lineup condition, where only the target or replacement was wearing clothing similar to that worn during the crime. Lindsay et al. found that the rate of correct identifications (target-present conditions) was not significantly influenced by clothing condition, although the greatest number of correct identifications occurred in the biased condition and the least in the dressed alike conditions. Important differences did emerge in the target-absent condition, however, with significant differences among rates of false identification according to clothing condition. The highest rate of selections of the replacement occurred in the biased condition, suggesting that participants' selections were influenced by the clothing worn by lineup members.

The issue of clothing bias has to date not been studied with children in an eyewitness context. This issue is important due to the recent dramatic increase in legal cases involving children as eyewitnesses (Bala, Lee, Lindsay, & Talwar, 2000; Bruck, Ceci, & Hembrooke, 1998; Gray, 1993). Although most cases that involve child witnesses concern crimes committed by a close relative or known trusted adult, many cases also involve strangers, for example child abduction and molestation (Lindsay, Pozzulo, Craig, Lee, & Corber, 1997; Pozzulo & Lindsay, 1998, 1999). Because children in these cases often are the only eyewitnesses, it is vital that lineups are used that aid children in identifying guilty individuals while minimizing the chance that an innocent party is selected. The role of clothing cues is one issue that must be empirically addressed in order to advance the development of such lineup procedures.

Although there is no direct empirical evidence to substantiate the existence of a clothing bias in children's lineup identifications, a comparable effect has been documented in the developmental face recognition literature. Specifically, there is evidence of *paraphernalia* effects, in which the manipulation of an item such as a hat, scarf, or glasses affects children's judgments of identity (e.g., Baenninger, 1994; Carey & Diamond, 1977; Flin, 1985; Freire & Lee, 2001). For example, Carey and Diamond (1977) presented 6- to 10-year-olds with a series of forced-choice trials. In each trial, the child first saw a picture of an individual wearing an item of paraphernalia such as a hat or pair of glasses. After a short delay the child was given the task of identifying the target from a pair of photographs. One photograph was of the target, now without the paraphernalia item, and the other was of a person who bore a general facial resemblance to the target and was wearing the paraphernalia previously worn by the target. Carey and Diamond found that younger children were more likely than were older children to choose the photo of the person wearing the paraphernalia item in the recognition phase, and thus less likely to correctly choose the photo of the target.

Effects of paraphernalia also are dependent on whether they serve to discriminate between a target and distractors (Baenninger, 1994; Freire & Lee, 2001). For example, Freire and Lee (2001) first trained children to distinguish a target face from three similar faces until they could do so reliably. Next, they completed two types of trials that included paraphernalia manipulations. In *uniform* trials a hat was added to photos of the target and the three distractors. In *mixed* trials a hat was added to the target and one of the distractors while the other two distractors remained bareheaded. Freire and Lee found that identification of the target was less accurate in mixed trials, as children tended to pick one of the two distractors without the hat (whose general

appearance is similar to the target seen earlier). This finding suggests that children were making some selections on the basis of overall similarity of the images, that is, they were using a relative judgment strategy. As in Carey and Diamond's study there was an inverse relationship between susceptibility to the hat manipulation and age. Together, these results suggest that clothing-biased lineup procedures may lead children to make incorrect identifications, similar to those made by adults in Lindsay et al. (1987).

However, this suggestion is premature. An important aspect of face recognition studies demonstrating paraphernalia effects in children is that paraphernalia items such as a hat, bandana, or glasses, cover part of the face (e.g., Baenninger, 1994; Carey & Diamond, 1977; Freire & Lee, 2001). As such, they may be interfering with children's ability to process the face itself. For example, with a hat or bandana, the external contour of a face is changed, as is the relationship between major face features and the external contour. Children may have difficulty in extracting the information necessary to gauge the identity of the face. Alternatively, children may not have difficulty in extracting such information. Rather, they are misled by the clothing cues, resulting in the paraphernalia effect. In typical paraphernalia studies, the contributions of these two factors are often confounded. To address this confound, one needs to use paraphernalia items that do not interfere with processing of the face itself, such as the style or color of clothing worn by lineup members (Lindsay et al., 1987). More importantly, in actual police lineups, it is highly unlikely that a suspect would wear a hat or bandana, but relatively common that a suspect's clothing may differ from that of foils. It is also possible that in a lineup the suspect may wear clothes that are similar to those worn during the crime. This may occur because the suspect habitually wears the same style of clothing and/or was arrested in part because his or her clothing matched a description provided by witnesses. Thus, it is more legally relevant to examine the impact on children's lineup selections biased by clothing, paraphernalia that does not obscure the face itself.

This study adapted the procedure of Lindsay et al. (1987) to investigate whether clothing-biased lineup procedures have an impact on children's lineup identification accuracy. Four- to 14-year-olds saw a series of 12 video clips, each showing a different individual. After each clip they were presented with a simultaneous photo lineup in which the target was present or absent. Three different clothing manipulations were included. In the *same clothing* lineup condition, individuals in the photographs wore the same colored shirt as the target in the video. As in the Lindsay et al. study, this condition allowed for the possibility that the clothing could act as a cue to assist the witness to select the guilty party from a target-present lineup. In the *different clothing* lineup condition, individuals all wore the same colored shirt, but the color was different from the shirt worn by the target in the video clip. This lineup represents a clearly unbiased lineup and one that includes no cues associated with the target or replacement. In the *biased clothing* lineup condition the target (in the target-present condition), or replacement (in the target-absent condition), was shown in the same colored shirt as the target in the video clip. Remaining lineup members wore shirts different in color from those of both the target and each other.

On the basis of the existing, albeit limited research, a number of predictions were made. With regard to target-present trials, Lindsay et al. (1987) expected but

found no effect of clothing condition on the number of correct identifications with adult participants. However, results of paraphernalia studies suggest that children may rely on clothing cues in their identifications. Thus, we expected an effect of clothing condition in target-present trials, with a greater number of correct identifications occurring in the biased condition. With regard to target-absent trials, adult participants in Lindsay et al. (1987) selected the replacement significantly more often in the biased clothing condition than in either of the two conditions in which the replacement did not stand out because of his clothes. Paraphernalia effects in the developmental literature also lead to the suggestion that children's selections will be influenced by clothing cues (e.g., Carey & Diamond, 1977; Freire & Lee, 2001). Thus, in the present study, more false identifications were expected in the biased target-absent condition than in either the same clothing or different clothing conditions. Further, because there is a decrease in vulnerability to paraphernalia effects as age increases, it was expected that selections of the replacement would decrease with age. Finally, consistent with findings in the eyewitness literature, correct rejections were expected to increase with age (Pozzulo & Lindsay, 1998).

METHOD

Participants

Children (120 boys, 108 girls) between the ages of 4 and 14 years participated. They lived in regions ranging from a small city to a large metropolitan area. A summary of the gender and age of participants appears in Table 1. Written parental consent was obtained or consent was granted from the head of school for the children's participation in the study. Verbal assent was received from the children.

Materials

Models and Clothing

Six male and six female university students of European descent served as target individuals. Approximately 150 additional male and female students of a similar age

Table 1. Summary of Age and Gender of Participants

Age(years)	*n*	Male/female	*M* (years; months)	*SD* (months)
4	21	7/14	4; 8	3
5	35	22/13	5; 6	3
6	33	17/16	6; 6	3
7	35	19/16	7; 7	3
8	22	10/12	8; 6	4
9	10	5/5	9; 6	4
10	21	8/13	10; 6	3
11	22	15/7	11; 6	3
12	14	9/5	12; 6	3
13	14	7/7	13; 5	3
14	1	1/0	14; 4	—

and ethnic background served as the pool from which lineup foils were selected. Six T-shirts of clearly distinguishable colors were used in lineup manipulations (black, gray, blue, red, orange, and green).

Videotapes

A 20 s video clip was filmed of each target individual. Each target wore one of the six T-shirts. Shirt color was assigned such that each color was worn once by a male target and once by a female target. In each clip, the target said a few words of greeting, said his or her name, told a children's joke and then said a short goodbye. The following is an example of the jokes: "What do cats eat at parties? Mice cream!"

Stimulus videotapes contained all 12 video clips. To reduce interference between trials, clips of male and female targets were presented in alternating order. To allow the experimenter time to pause the video between clips, a screen that presented information about the tape and clip number appeared between clips. Six different stimulus tapes that differed only in the presentation order of the 12 clips were created. One was randomly selected for use with each child, so that the order of presentation of the clips, and therefore of the models and shirt colors, differed across participants.

Photographs

Each target individual was photographed once while wearing the shirt he or she wore in the video and once while wearing one of the other colored shirts; shirt assignment for the second set of photographs was done such that each color was worn once by a male target and once by a female target. Each individual in the pool of foils was photographed six times, once while wearing each T-shirt. Six of these individuals were then selected for each target, to complete lineups. Selection of foils was based on general facial resemblance to the target and similarity of length and color of hair. In all pictures, individuals posed with a neutral facial expression and were photographed from the chest up, ensuring that each person's face and the color of the shirt he or she was wearing were clearly visible. All photographs were taken under similar lighting conditions using a digital camera, against a uniform background, and from a distance of approximately 5 ft. Photographs required for the final set of lineups (see below) were each 11.5-cm wide $\times$9.5-cm high. They were printed on photographic paper on a high-quality color laser printer and laminated.

Lineup Construction

Pictures of the target and five of the six foils comprised lineups in the target-present conditions. A photo of the sixth foil, the *replacement*, was substituted for that of the target in the target-absent conditions. A haphazard selection determined which one of the six foils would be omitted from the target-present lineup and serve as the replacement in the target-absent lineup; that is, no attempt was made to select the most similar foil as the replacement.

Pictures were then selected to complete three different types of lineups for each target. For the same clothing condition, pictures were selected in which lineup members, including the target or replacement, wore the same colored shirt as the

target wore in the video clip. For the different clothing condition, pictures were selected so that all lineup members, including the target or replacement, wore the same colored shirt, but this color was different from that worn by the target in the video clip. For the biased condition, photographs were chosen such that only the target in the target-present condition, or the replacement in the target-absent condition, was wearing the same colored shirt as worn by the target in the video clip. Photographs of the five remaining lineup members were chosen so that each was wearing a shirt different in color from the target, or replacement, and from each other.

Procedure

Children were tested individually at their school in a quiet location. Each child was told that he or she would watch a video of people telling jokes. The child was told that after each joke he or she would be shown some photographs and would be asked to pick the photo of the person who had just appeared on the television telling the joke. Further, the child was told that sometimes he or she would see the person's picture, and that sometimes the person's picture would be missing. It was explained to the child that if the correct picture was there, he or she would need to point it out. If the correct picture was not there, he or she would just say the picture was not there. The child was asked if he or she had any questions. Once any questions were answered the video was shown on a combination color television–video cassette recorder with a 14-in. screen.

The tape was paused after the first clip was viewed. The experimenter shuffled the six photographs comprising the lineup, and placed them on a table or desk in front of the child in a random order of two rows of three, one above the other. The child was told to examine all of the pictures carefully and was reminded that a picture of the person who just told the joke may or may not be among the pictures. The instruction was repeated to point out the correct picture to the experimenter if it was there, and to say the picture was not there if none of the pictures were the correct one. After the first selection was made this procedure was repeated for the subsequent 11 trials, including a reminder each time that the target may or may not be present in the array of pictures. The child was given a break after the first six clips were shown and was given a small prize. The child was asked whether he or she would like to continue to view six more clips to earn another prize.

The 12 lineups presented to each child generated a 2 (sex of target) $\times$ 2 (target-present vs. absent) $\times$ 3 (clothing condition: same, different, and biased) within-participants design. There were never more than two consecutive lineups of either the target-absent or target-present type. Lineups were also presented such that the first six trials included one target-present and one target-absent lineup for each clothing condition, as did the last six trials. Because lineups were unique for each target individual and each target was shown only once in the stimulus videotapes, children did not see more than one photograph of any given individual (target or foil) throughout the session. Across all participants, each of the 12 target individuals served as a target approximately an equal number of times in each of the six lineup conditions.

RESULTS

A correct response was scored in the target-present condition when the target was selected from the lineup, whereas a correct response in the target-absent condition occurred when the lineup was rejected. Because correct responses are not the same in the target-present and target-absent conditions the corresponding data were analyzed separately. There were six lineups of the target-present and target-absent conditions, two lineups of each of the clothing conditions (same, different, and biased clothing). Thus, there were two same clothing/target-present trials, two different clothing/target-present trials, and so forth, allowing a score out of two to be calculated for each lineup type. A score of 0 represents no correct responses, 1 represents one correct response, and 2 represents two correct responses. The scores out of two were then converted into proportions.

Results of three children were excluded from all analyses because they only partially completed the study and provided insufficient data to include in any of the analyses. An additional seven participants were missing data for a single trial. The reason for two missing trials was that the child responded "I don't know," rather than either picking someone out of the lineup or rejecting it. There were missing data for five other trials because of experimental error. The seven children with missing data for a single trial were excluded from analyses dependent on the missing trial but included in all unaffected analyses.

Preliminary analyses found no effects of sex of either the child or target on any of the response measures. Therefore data were collapsed across boys and girls, and across sex of target, for all ensuing analyses.

Data for target-present and target-absent lineups were each analyzed in two ways. Responses on target-present trials were analyzed in terms of (1) proportion of correct identifications; and (2) proportion of incorrectly rejected lineups. The second analysis addressed the issue of whether clothing condition and age influenced if someone was chosen from the lineup (either correctly or incorrectly), or the lineup was incorrectly rejected. Responses for target-absent lineups were analyzed in terms of (1) proportion of correct rejections (i.e., correctly reporting that the target was not in the lineup); and (2) proportion of incorrect selections of the replacement individual.

Children's performance in terms of the above measures was analyzed in four separate general linear models. There were three terms in each model: *clothing condition*, a categorical, within-participants variable with three levels that was dummy coded for purposes of these analyses; *age* (in months), a continuous between-participants variable; and an interaction term (Clothing condition $\times$ Age). All means reported are estimated marginal means. It should be noted that because the dependent variable was a proportion in each case, analyses were redone using an arcsine transformation. All analyses of the transformed data produced the same effects as did the proportion data. The proportion data are presented here for clarity.

Target-Present Trials

Figure 1 presents data for proportion correct and proportion of rejected lineups in the target-present trials, for each clothing condition and across all ages. Figure 2

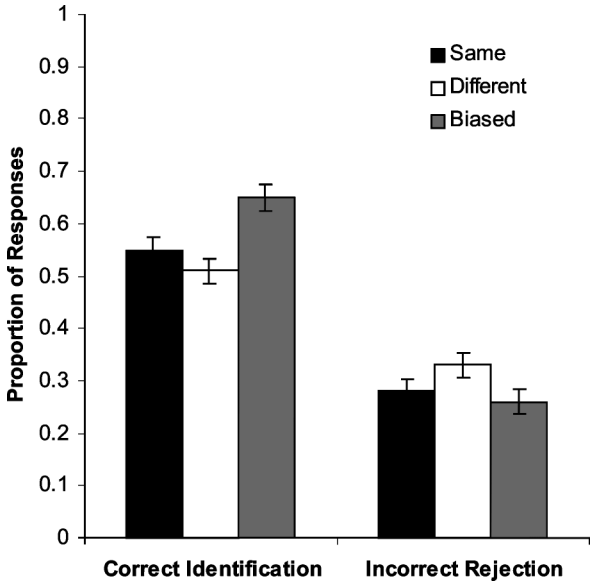

Fig. 1. Proportion of correct identifications and of lineup rejections (+*SE*) in target-present trials for the three clothing conditions.

shows the same two measures as a function of age, collapsing across clothing condition. For purposes of illustration children are grouped by age, in years, with the lone 14-year-old included with the 13-year-olds. The same is true of the corresponding figure of performance in target-absent trials below. The effect of clothing condition on proportion of correct identifications was significant, $F(2, 442) = 10.32$, $p < .01$ (partial $\eta^2 = .05$). Planned contrasts revealed a significant difference between the biased ($M = 0.65$) and same ($M = 0.55$) clothing conditions, $F(1, 221) = 10.48$, $p < .01$, and

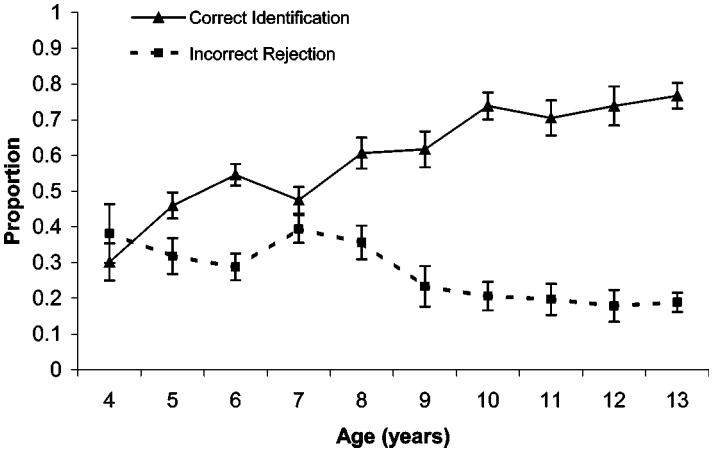

Fig. 2. Proportion of correct identifications and of lineup rejections (+*SE*) in target-present trials as a function of age.

between the biased and different ($M = 0.51$) clothing conditions, $F(1, 221) = 19.40$, $p < .01$. The difference between the same and different clothing conditions was not significant. Thus, more selections of the target occurred in the biased condition than in either the same or different clothing conditions. There was also a significant effect of age, with older children making more correct identifications, $F(1, 221) = 87.66$, $p < .01$ (partial $\eta^2 = .28$). The interaction between clothing condition and age was not significant, $F(2, 442) = 2.21$, *ns* (partial $\eta^2 = .01$).

The effect of clothing condition on proportion of incorrectly rejected lineups was also significant, $F(2, 442) = 4.25$, $p < .05$ (partial $\eta^2 = .02$). There were no specific predictions made with respect to the effect of clothing condition on rejection of target-present lineups. A posthoc analysis with Bonferroni correction for multiple comparisons indicated a significant difference between the biased ($M = 0.26$) and different clothing ($M = 0.33$) conditions in terms of proportion of rejected lineups, $F(1, 221) = 7.75$, $p < .01$, with other differences not significant ($M = 0.28$ for same clothing condition). Thus, biased lineups were rejected less frequently than were different clothing lineups. There was also a significant effect of age on rejection of target-present lineups, with older children rejecting fewer lineups, $F(1, 221) = 14.08$, $p < .01$ (partial $\eta^2 = .06$). The interaction between clothing condition and age was not significant, $F(2, 442) = .08$, *ns* (partial $\eta^2 = .00$).

To examine whether our overall results could be attributed to the repeated measures design of the study, we carried out separate analyses on the first six trials only, using Cochran's Q test. These trials represented children's first exposure to each of the six lineup combinations (3 clothing conditions × presence/absence of target). Any effects present in these six trials would therefore suggest that children began the study with biases already in place, or else were very susceptible to their development. With respect to correct choices in the three clothing conditions, 174 children were correct in the biased trial, 145 in the same trial, and 110 in the different condition, $Q(2) = 38.39$, $p < .01$. Thus, equivalent to the overall pattern of results, children in the first half of trials were more accurate in the biased target-present condition relative to the same and different conditions. In terms of incorrectly rejecting target-present lineups, significant differences observed over all trials were also present when looking at only the first six trials, $Q(2) = 22.95$, $p < .01$. As in the overall case, the fewest rejections occurred with biased lineups and the greatest number with different lineups (44, 52, and 83 rejected lineups in biased, same, and different clothing conditions, respectively).

Target-Absent Trials

Figure 3 presents data for proportion of correct rejections and proportion of selections of the target replacement for the three clothing conditions, collapsed across age. Figure 4 shows these measures as a function of age, collapsed across clothing conditions. The effect of clothing condition on proportion of correct rejections was not significant ($M = 0.62$, 0.63, and 0.60 for same, different, and biased clothing conditions, respectively), $F(2, 438) = .24$, *ns* (partial $\eta^2 = .00$). There was a significant effect of age, with younger children making fewer correct rejections, $F(1, 219) = 37.91$,

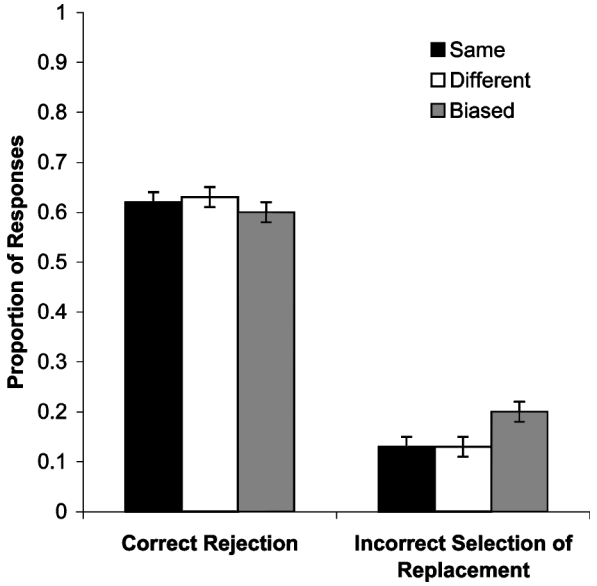

Fig. 3. Proportion of correct rejections and of selections of the target replacement (+*SE*) in target-absent trials for the three clothing conditions.

$p < .01$ (partial $\eta^2 = .15$). The interaction between clothing condition and age was not significant, $F(2, 438) = 1.23$, *ns* (partial $\eta^2 = .01$).

The effect of clothing condition on incorrect selection of the replacement was significant, $F(2, 440) = 6.15$, $p < .01$ (partial $\eta^2 = .03$). Planned contrasts revealed a significant difference between the biased ($M = 0.20$) and same ($M = 0.13$) clothing conditions, $F(1, 220) = 6.97$, $p < .01$, and between the biased and different ($M = 0.13$)

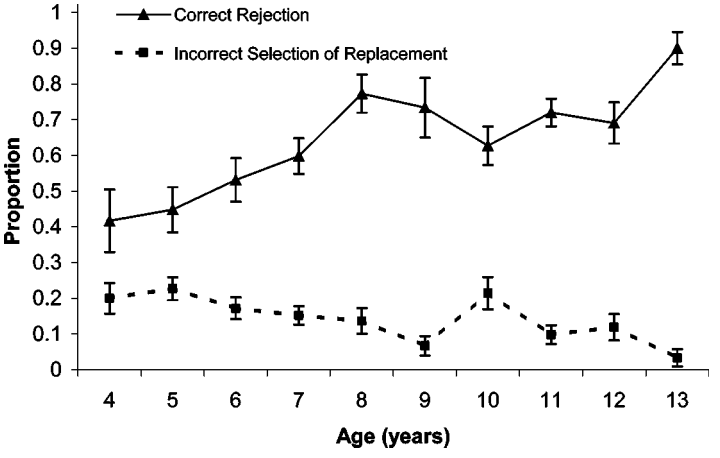

Fig. 4. Proportion of correct rejections and of selections of the target replacement (+*SE*) in target-absent trials as a function of age.

clothing conditions, $F(1, 220) = 9.63$, $p < .01$. Thus, more incorrect selections of the replacement occurred in the biased condition than in either the same or different clothing conditions. The effect of age on proportion of selections of the replacement also was significant, $F(1, 220) = 13.62$, $p < .01$ (partial $\eta^2 = .06$), with older children incorrectly choosing the replacement individual from target-absent lineups less frequently than younger children. The interaction between clothing condition and age was not significant, $F(2, 440) = .04$, *ns* (partial $\eta^2 = .00$).

As with target-present trials analyses were done separately on the first six trials, to examine whether the pattern observed over all trials held. Similar to the overall analysis, analysis of the first six trials indicated no effect of clothing condition on proportion of correct rejections, $Q(2) = 2.44$, *ns* (number of correct rejections were 127, 139, and 139 for biased, same, and different conditions, respectively). Incorrect selections of the replacement individual from target-absent lineups in the first six trials were significantly different for the three clothing conditions, $Q(2) = 22.22$, $p < .01$. Approximately twice as many children chose the replacement in the biased condition as in either the same or different condition (60, 26, and 31 children, respectively).

DISCUSSION

Lineup decisions by children in this study were influenced by the clothing worn by lineup members. In target-present trials the target was chosen more frequently in the biased clothing condition than in either the same clothing or different clothing conditions. Similarly, in target-absent trials the replacement was incorrectly selected most frequently in the biased condition. Thus whereas a shirt color that matched that seen in the video clip assisted identification of the target when the target was present in the lineup, it also led to more false identifications when the target was absent.

The finding that biased clothing cues led children to make false identifications when the target was absent from the lineup is consistent with reports of paraphernalia effects in the face recognition literature (e.g., Baenninger, 1994; Flin, 1985; Freire & Lee, 2001) and with a previous eyewitness study with adult participants (Lindsay et al., 1987). All of these studies provide evidence of misidentifications made on the basis of visual cues unrelated to facial identity. What is striking about the present findings is the nature of the cue. Paraphernalia effects have been demonstrated for items that occlude part of the face, for example a hat or glasses. Such effects may be due in part to a disruption to face processing, in addition to providing a misleading cue to identity. Lindsay et al.'s biased clothing condition involved a university sweatshirt (Lindsay et al., 1987), a distinctive piece of clothing with a large image of drunken bisons (the name of a university's athletic teams). In contrast, in the present study what distinguished the replacement from the remaining foils in the target-absent biased condition was shirt color. This finding suggests that a distinguishing clothing cue, one that matches the one previously worn by a target, need not be distinctive in design to lead to misidentifications by young children.

The above finding in the target-absent trials also suggests that when children encode faces, the information encoded is in fact not limited to the face. Rather, they appear to encode a broader range of information, including the color of clothing. The

suggestion that children encoded information about shirt color in this study also is supported by the results of target-present trials. In these trials, the biased condition resulted in more correct selections of the target than either of the other two clothing conditions.

However, it is important to emphasize that children did not appear to use shirt color as the sole, nor primary, basis for identification. Indeed, although all clothing effects reported appear reliable, effect sizes were modest. The fact that shirt color was not the primary basis for identification is evident in the finding that none of our performance measures indicated significant differences between the same and different clothing conditions. This finding suggests that, although children encoded shirt color, a match in shirt color to that seen in the video clip was not a prerequisite for picking a lineup member, nor was a mismatch sufficient grounds for rejecting a lineup. Instead, the picture that emerges from our results is that children's decisions were based primarily on memory for faces, with shirt color playing a supplemental role in their lineup selections.

Findings related to lineup rejections also are of interest. Younger children are known to be reluctant to reject lineups relative to older children (e.g., Lindsay et al., 1997). This trend held in the present study for target-absent lineups, with older children more inclined to reject correctly a target-absent lineup. Younger children did, however, reject more lineups than older children in the target-present conditions. This result appears to contradict the notion that younger children are, in general, less likely to reject lineups. However, because older children were more accurate in target-present trials, there were more opportunities for younger children to reject a target-present lineup in absolute terms. To examine the issue of rejection rates of target-present lineups properly, rejections have to be considered as a percentage of total errors (i.e., rejections or picking someone other than the target). The rate of rejection of target-present lineups expressed as a percentage of total errors for 4- and 5-year-olds, 6- and 7-year-olds, 8- and 9-year-olds, 10- and 11-year-olds, and 12- to 14-year-olds was 57, 71, 81, 72, and 74% respectively. Thus, younger children do not reject more target-present lineups when rejection rates are adjusted to account for differing degrees of accuracy.

Age effects were evident in all aspects of performance. When the target was present, older children correctly identified the target more frequently than did younger children and rejected fewer lineups outright. Older children more accurately rejected target-absent lineups than did younger children, and less frequently incorrectly selected the replacement. Despite these age differences the factors of age and clothing condition did not interact, as would be predicted from studies of paraphernalia effects, which show younger children to be more vulnerable to such cues than older children (e.g., Baenninger, 1994; Freire & Lee, 2001). One explanation for this discrepancy is that shirt color is simply not a salient enough cue to produce paraphernalia effects that differ with age, which have typically been found with items such as distinctive hats or glasses. One of the mechanisms that may contribute to paraphernalia effects in the face recognition literature is that the salience of the paraphernalia item distracts the child, decreasing attention paid to the face, with this effect more pronounced for younger than older children (e.g., Freire & Lee, 2001). It is possible that, in the present study, although shirt color contributed to children's

choices from lineups, it nevertheless was not salient enough to reduce children's attention to faces to the extent that an unusual hat or pair of glasses might. Another possibility is that, because a hat or glasses cover part of the face, they interfere with encoding and recognition, and that this interference represents more of a hindrance to younger children than to older children. Finally, the existence of a clothing effect with adults (Lindsay et al., 1987) may indicate that clothing bias exists at all ages.

The design of this study contained elements of both face recognition and eyewitness tasks. Children were instructed that they were to remember the target individuals, and completed multiple trials. These aspects are typical of face recognition tasks but not eyewitness paradigms. In contrast, initial exposure to the targets in this study was via a dynamic image and children made decisions about individuals presented in lineups, typical of eyewitness tasks but not face recognition ones. Given the hybrid paradigm, it is noteworthy that the major results were consistent with predictions; that is, that biased clothing conditions resulted in more accurate identifications in the target-present trials and more misidentifications in the target-absent trials. These results also are consistent with findings of studies in which paraphernalia directly overlap with target and distractor faces. Such consistency of findings despite differences in methodology speaks to the robustness of the phenomena of paraphernalia/clothing bias effects. It also lends support to the suggestion that nonface related information does influence young children's memory and identification of faces.

The robustness of our results is also evident in that clothing effects were present even in the first six trials, when children were seeing each type of lineup for the first time. As pointed out by an anonymous reviewer, despite the fact that children were not explicitly told about the clothing manipulations, they could have formed hypotheses about the relevance of shirt color on the basis of the early trials. However, analyses restricted to the first 6 trials indicated that all the same effects were present as described for the full 12 trials. These findings do not rule out the possibility that children's performance in later trials was influenced by what they learned in preceding trials, or by any hypotheses they may have formed about the role of shirt color. They do, however, suggest that children either entered the experiment with biases about clothing cues, or else were highly susceptible to the development of clothing biases. Future research can examine these alternatives more directly.

From an applied perspective the present results highlight the importance of conducting lineups that are unbiased with respect to clothing cues, given the higher rate of false identifications in the target-absent, biased condition. Thus, fair lineups in terms of clothing are clearly preferred options when adult or child witnesses attempt to identify suspects. It should be noted, however, that children's performance in the two nonbiased clothing conditions was similar on all measures. This finding suggests that although clothing cues are encoded by children in their memory for faces, it is unnecessary that the clothing worn by lineup members be the same as that worn by the suspect at the scene of the crime. This suggestion has the practical advantage that lineup members can be dressed in generic clothing such as sweatshirts of a common color. Special efforts need not be made to locate clothing similar to those worn during the crime.

This study represents a first step to understanding the role of clothing cues in children's identification of individuals from lineups. Additional studies are needed

to clarify a number of critical issues regarding clothing bias. For example, one issue is whether effects hold when children's memory for the target face is incidental, as in more traditional eyewitness paradigms. In this study children were told to remember the targets, and were exposed to a head-on view of each. Another issue was mentioned above, that of the effects of more distinctive clothing. It is possible that more distinctive cues than a different color of shirt (e.g., an unusual looking jacket or brightly colored hat), would lead to a more pronounced bias. The emotional context of the viewing condition is another factor that could influence whether bias is observed, or its extent. Children in this study watched individuals tell jokes rather than watching a negative event such as a staged theft, as is typical in an eyewitness paradigm.

The specific instructions provided to an eyewitness can have a significant influence on his or her accuracy in a lineup task (see Steblay, 1997, for a review). An especially interesting question raised by our study relates to how results might differ if children were informed that they should ignore clothing as a cue in making their lineup decisions. Would they be able to use this information to focus their efforts on processing of the target faces, or would there be little effect on their performance? This manipulation would provide some insight into whether children's susceptibility to clothing bias is relatively automatic on the one hand, or whether it is a result of more deliberate processing. If the clothing bias observed in the present study is relatively automatic, telling children to ignore clothing cues should have little bearing on performance. In contrast, if the clothing bias is the result of relatively deliberate processing, then we might expect that telling children to ignore clothing cues may have differing effects on the performance of younger and older children. Specifically, such an instruction might negate, or at least reduce, the bias for the older children. Younger children, because they may have more difficulty exerting deliberate control over their perceptual and decision-making processes, may experience more difficulty in ignoring the clothing cues and thereby might continue to show evidence of clothing bias. Regardless of the effects such instructions may have, police procedures do not provide such instructions. As a result, the data as collected reflect expected patterns of eyewitness accuracy given current police practice.

Finally, the influence of different ways of conducting lineups also is worthy of investigation. In recent years, alternative procedures have been developed that are better suited to children than the traditional simultaneous lineup used in this study. In an *elimination* lineup, children are first presented with a simultaneous lineup and are asked to indicate which person looks most like the target. Having done so, the remaining pictures are removed and the child is asked whether the chosen picture shows the correct person or merely someone resembling the correct person. Research illustrates that although the elimination lineup produces equivalent rates of correct identifications to a simultaneous lineup, rates of false identifications by children are significantly reduced (Pozzulo & Lindsay, 1999). Whether this advantage is maintained when the lineup is biased in terms of clothing is of interest. If it is shown to be resistant to clothing bias the elimination lineup becomes an even more attractive alternative to the traditional simultaneous lineup.

In summary, results of the present and these future studies should provide a foundation for a better understanding of the effects of clothing cues in general, and clothing bias in particular, on children's identifications from lineups. In turn, this

knowledge would allow for the development of improved procedures for carrying out lineup identifications with young children.

ACKNOWLEDGMENTS

This study was supported by a grant from the Social Science and Humanities Research Council of Canada to the second author. The authors thank children and their schools for participating in this study. The authors' gratitude is also extended to the anonymous reviewers and the editor for their constructive comments on earlier versions of this paper.

REFERENCES

Baenninger, M. (1994). The development of face recognition: Featural or configurational processing? *Journal of Experimental Child Psychology*, *57,* 377–396.

Bala, N., Lee, K., Lindsay, R. C. L., & Talwar, V. (2000). A legal and psychological critique of the present approach to the assessment of the competence of child witnesses. *Osgoode Hall Law Journal*, *38,* 409–452.

Bruck, M., Ceci, S. J., & Hembrooke, H. (1998). Reliability and credibility of young children's reports. *American Psychologist*, *53,* 136–151.

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

Dysart, J. E., Lindsay, R. C. L., MacDonald, T. K., & Wicke, C. (2002) The intoxicated witness: Effects of alcohol on identification accuracy from show-ups. *Journal of Applied Psychology*, *87,* 170–175.

Flin, R. H. (1985). Development of face recognition: An encoding switch? *British Journal of Psychology*, *76,* 123–134.

Freire, A., & Lee, K. (2001). Face recognition in 4- to 7-year olds: Processing of configural, featural, and paraphernalia information. *Journal of Experimental Child Psychology*, *80,* 347–371.

Gray, E. (1993). *Unequal justice: The prosecution of child sexual abuse*. New York: MacMillan.

Lindsay, R. C. L., Nosworthy, G. J., Martin, R., & Martynuck, C. (1994). Using mug shots to find suspects. *Journal of Applied Psychology*, *79,* 121–130.

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.

Lindsay, R. C. L., Wallbridge, H., & Drennan, D. (1987). Do the clothes make the man? An exploration of the effect of lineup attire on eyewitness identification accuracy. *Canadian Journal of Behavioural Science*, *19,* 464–478.

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.

Lindsay, R. C. L., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology*, *70,* 556–564.

Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology*, *66,* 482–489.

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

Pozzulo, J. D., & Lindsay, R. C. L. (1999). Elimination lineups: An improved identification procedure for child witnesses. *Journal of Applied Psychology*, *84,* 167–176.

Steblay, N. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior*, *21,* 283–297.

Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology*, *14,* 89–103.

Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., & Fulero, S. M. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist*, *55,* 581–598.

Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distractors for eyewitness lineups. *Journal of Applied Psychology*, *78,* 835–844.