



## The role of experimenter familiarity in children's eyewitness identification

Lesley Calderwood , Carrie Ballantyne  and Kimberley Slee

*Division of Psychology, University of the West of Scotland, Paisley, UK*

Child eyewitnesses show a high false identification rate on target-absent (TA) lineups despite good performance on target-present (TP) lineups. One explanation is that children feel a social pressure to choose when presented with a TA lineup. We investigated whether experimenter familiarity would reduce social pressure and improve accuracy on TA lineups. Children (5–7 years,  $N=120$ ) watched a short video of a staged theft; 1–2 days later they completed a TP or TA lineup with a familiar or unfamiliar experimenter. Experimenter familiarity had an impact on lineup response in TA lineups only, with more correct 'not there' and fewer 'not sure' responses when the children were familiar with the experimenter. The results provide further evidence to support the social aspect of eyewitness identification decisions in children and provide a possible strategy to improve identification accuracy for those working with children in the criminal justice system.

**Keywords:** child eyewitness memory; choosing behaviour; eyewitness identification procedures; familiarity of lineup administrator; improving identification accuracy; lineup options; social pressure.

**Article History:** Received 5 September 2022; Accepted 25 January 2023

Children's ability to act as a reliable eyewitness has long been an area of scientific debate. It is estimated that around a third of witnesses are children (Memon et al., 2011), and it is therefore imperative that we understand the factors underlying reliable eyewitness identification in children. One consistent finding from the eyewitness literature is that children are as accurate as adults when identifying a suspect on a target-present (TP) lineup from five years of age (Pozzulo & Lindsay, 1998), with only some studies reporting age-related improvements on TP lineups (Fitzgerald & Price, 2015; Keast et al., 2007). In contrast, a large body of evidence shows a lack of age-related improvement on target-absent (TA) lineups with children showing a strong tendency to make a false identification (for example, Beal

et al., 1995; Fitzgerald & Price, 2015; R. C. Lindsay et al., 1997; Parker & Carranza, 1989; Parker & Ryan, 1993; Pozzulo & Balfour, 2006). The same pattern has also been reported using video as opposed to photo lineups (Havard et al., 2012). It is therefore critically important to identify the factors that influence children's eyewitness identification decisions given the real-world implications of erroneously choosing someone when the target is not present in the lineup. The present study sought to examine the effects of two potentially important factors – familiarity with the lineup administrator and options for rejecting the lineup. We investigated whether young children would be more likely to correctly reject a TA lineup with a familiar administrator and

Correspondence: Lesley Calderwood, Psychology, University of the West of Scotland School of Education and Social Sciences, Paisley, PA1 2BE, UK. Email: [Lesley.Calderwood@uws.ac.uk](mailto:Lesley.Calderwood@uws.ac.uk)

when more explicit options to reject the lineup were available.

One of the most likely explanations for the contrasting pattern of performance on TP and TA lineups by children relates to the social factors involved when children are presented with a TA lineup (Brewer & Day, 2005; Havard, 2014; Memon et al., 2003; Pozzulo et al., 2012). It is argued that the act of presenting a lineup to an eyewitness indirectly suggests that the target must be present and that, as a result, child witnesses feel a strong social pressure to choose someone from the lineup (Beal et al., 1995; Ceci et al., 1987; Raskin & Yuille, 1989). One aim of the current study was therefore to examine whether social pressure to choose would be reduced when children were familiar with the person carrying out the lineup procedure. If false identifications on TA lineups are reduced when the lineup is administered by a familiar person, then this result would lend further support to the role of social factors in children's eyewitness identification errors.

### ***Role of familiarity***

Previous research has shown that the behaviour and characteristics of the interviewer can influence the accuracy of children's verbal recall of an event. Evidence suggests that young children are more accurate when interviewed by their parents than by a stranger (Flvush et al., 1991) and are less susceptible to misleading questions when they are interviewed by someone familiar to them (Jackson & Crockenberg, 1998; Quas & Schaaf, 2002). Bjorklund et al. (2000) found that 5–7-year-olds (the age group used in the current study) were more accurate when questioned a second time by the same interviewer. However, Ricci et al. (1996) found that 5-year-olds were less accurate when questioned by their parent than by an unfamiliar adult. Other researchers argue that interviewer *support* plays a more important role in obtaining accurate information from child witnesses than interviewer familiarity and that children are less suggestible when

interviewed by someone supportive (Brubacher et al., 2019; Davis & Bottoms, 2002). A supportive interviewer may therefore help to reduce the implicit suggestibility experienced by children when presented with a TA lineup and subsequently increase correct rejections.

Evidence that children's identification decisions can be influenced by the social cues and behaviour of the interviewer was found when children aged 9–10 years made more false identifications when the lineup administrator wore a police-like uniform than when they wore casual clothing, while there was no such effect on correct identifications from TP lineups (Lowenstein et al., 2010). Rush et al. (2014) found that having a supportive interviewer (someone who was dressed casually, introduced themselves, smiled and gave positive feedback) reduced false identifications in 7–8-year-olds and 12–14-year-olds who had experienced high stress levels at the time of encoding with no effect of interviewer on correct identifications in TP lineups. Taken together, these results suggest that the perception of the lineup administrator as someone of authority may lead to more false identifications and that children may be more willing to indicate that they do not see the perpetrator in the lineup to someone who is familiar to them.

To our knowledge, only one study has manipulated the familiarity of the lineup administrator. Ricci et al. (1996) found that 5-year-old children were less accurate and were more likely to change their answer when lineups were administered by their parent than by a stranger. The role of familiarity in improving children's eyewitness accuracy is therefore unclear, and one of the aims of the current study is to examine whether having the lineup administered by a familiar person will lead to an increase in correct rejections on TA lineups. If this is the case, then this finding would lend further support to the role of social factors when accounting for children's high error rate on TA lineups.

### *Lineup response options*

In addition to investigating the role of familiarity on lineup accuracy, the current study also explores the options for children to reject a lineup. One reason for the difference in accuracy between children and adults on TA lineups may be linked to the different task demands of TA and TP lineups. In a TP lineup, a child can point to their answer, whereas a verbal response is required to reject a TA lineup. Various studies have addressed this difference in response requirements by developing the TA lineup options to include a salient option for children to point to when the person is not there, such as a 'not here' card (Beal et al., 1995), a 'Mr Nobody' card (Davies et al., 1988) and more recently the 'wildcard', consisting of a silhouette with a question mark superimposed on it. The wildcard substantially increased correct rejection rates on TA lineups with both 8–11-year-olds (Zajac & Karageorge, 2009) and 5–7-year-olds (Karageorge & Zajac, 2011) compared to a standard TA lineup where children have to verbalise their 'not there' response. Similar beneficial effects have been found when a tree was included in the lineup for children to point to if they could not see the target (Dunlevy & Cherryman, 2013) and when a silhouetted mystery man was included in the video lineup for children to select when they could not see the target (Havard & Memon, 2013). It is therefore generally accepted that the inclusion of a salient option in the lineup that children can select to indicate that they do not see the target helps to reduce the task demands of TA lineups and also satisfies children's strong desire to choose someone in order to please the experimenter. It is important to note that while all of these techniques have improved children's ability to correctly reject a TA lineup, their good performance on TP lineups is maintained.

Recent research in this area has looked at combinations of these different techniques to further improve children's performance on TA lineups. For example, Thompson et al. (2020)

combined the elimination procedure with the wildcard and found that this method was the most effective in reducing false identifications. While the 'elimination with wildcard' procedure seems to be a promising technique to further improve children's identification accuracy on TA lineups, other researchers have argued that the wildcard may be interpreted more along the lines of a 'not sure' response rather than a definitive 'not there' response, and it has been suggested that the lineup should contain both 'not sure' and 'not there' options (Pica et al., 2020; Wells et al., 2020). An updated review of guidelines for eyewitness identification procedures in the United States has recommended that in addition to including an explicit statement that the target may or may not be present in the lineup, the lineup should also contain clearly marked 'not there' and 'do not know' options beneath the array of faces (Wells et al., 2020). One of the aims of the current study was therefore to explore how the inclusion of both 'not there' and 'not sure' response options would influence identification responses in child eyewitnesses.

A small number of studies have included both a 'don't know' and a 'not there' option in the lineup with child participants and have shown that child witnesses rarely use the 'don't know' option (Dekle et al., 1996; Memon & Rose, 2002; Pozzulo & Lindsay, 1997, with child witnesses aged 5–6, 8–9, and 10–14 years, respectively). Brewer et al. (2010) compared identification accuracy in 9–14 years across a range of conditions and found that the combination of the 'not sure' option with the instruction to be accurate seemed to be the most effective technique in reducing false identifications. Of interest to the current study was the finding that the inclusion of both 'not there' and 'not sure' responses in the lineup options made no difference to the overall number of false identifications compared to when there was only the 'not there' option (52% vs. 51%, respectively). There is therefore a clear need for more research to look at the impact on children's accuracy

when both ‘not there’ and ‘not sure’ options are presented within the lineup and to investigate the most effective format for the ‘not sure’ option. In the current study, the ‘not sure’ option was presented in the form of a ‘mystery boy’ (a silhouette with a question mark superimposed on it) alongside a ‘not there’ option.

### ***Confidence–accuracy relationship***

The value of eyewitness confidence ratings has been debated for many years in both the adult and child eyewitness literature with confidence generally regarded as an unreliable measure of identification accuracy (see Sporer et al., 1995, for a review). However, more recent approaches using calibration techniques suggest that confidence ratings may provide a more useful marker of accuracy than once thought (e.g. Juslin et al., 1996; Mickes, 2015; Olsson, 2000), and some researchers recommend that a clear statement of confidence is obtained from an eyewitness at the time of the identification decision (Valentine et al., 2007; Wells et al., 2020). It has also been suggested that there is a clearer relationship between confidence and accuracy for TP lineups whereas there is no relationship for TA lineups (Brewer & Palmer, 2010; D. S. Lindsay et al., 1998). There is some evidence of a similar pattern in child eyewitnesses where confidence levels were higher on TP lineups when they had made a correct identification (Havard & Memon, 2013; Havard et al., 2012). A further aim in the current study was therefore to investigate whether correct identifications on TP lineups were associated with higher confidence ratings and whether experimenter familiarity and lineup response options would influence these confidence ratings.

### ***The present study***

Given that children are often called upon to testify, it is imperative that we identify the most effective methods and procedures to ensure a reliable identification decision and

further explore the influence of social factors in their decision making. The current study compares lineup accuracy with a familiar versus unfamiliar experimenter while also adopting a lineup procedure including ‘not sure’ and ‘not there’ options. It is hypothesised that children will be just as accurate on TP lineups regardless of researcher familiarity, and that researcher familiarity will lead to more ‘not there’ responses on TA lineups. A further aim is to explore whether post-identification confidence ratings would be influenced by experimenter familiarity.

## **Method**

### ***Participants***

A total of 120 children aged between 5 and 7 years took part in the study. Due to the practical difficulties of working with this age group and the time required to develop familiarity with the experimenter, we were unable to adopt a true experimental manipulation of familiarity. As an alternative, we took the opportunity to investigate the effect of naturally acquired familiarity through the relationship of one of the researchers with children from a crèche facility where they worked. As a result, familiarity with the experimenter was a quasi-independent variable, but children were randomly assigned to the TP and TA lineup within each group.

Sixty children were recruited from a local Nuffield Health crèche facility and were familiar with the experimenter, and 60 children were recruited from a primary school and were unfamiliar with the experimenter. This number of participants is similar to previous studies that have used an eyewitness paradigm with the same age group (Havard & Memon, 2013; Karageorge & Zajac, 2011; Thompson et al., 2020). The mean age for familiar participants was 6.3 years of age, and the mean age for unfamiliar participants was 6.7 years of age. The familiar children had attended the crèche for a mean of 4.1 years (range = 1–5 years). The experimenter had worked at the

crèche facility for three years at the time of testing, and each child had known her for at least 6 months.

Ethical approval for this study was granted by the host institution. Written consent to carry out the study was granted by the crèche manager, the head teacher and legal guardians. Verbal assent was also obtained from each child on each day of testing. The children in each group were randomly assigned to the TP lineup or TA lineup.

### **Materials**

A recent review of 54 countries has highlighted wide variation in the guidelines for eyewitness identification procedures across the world (Fitzgerald et al., 2021). Photo lineups are used in the majority of cases in the United States (Fitzgerald et al., 2018), and while live lineups are the preferred medium in Australia and New Zealand (Evidence Act, 1995, 2006, respectively), the most prevalent form of identification in practice is a simultaneous photo lineup (Cullen et al., 2021; Fitzgerald et al., 2018). Fitzgerald et al. (2018) concluded that live lineups do not necessarily lead to higher accuracy, and given the practical difficulties in organising and administering a live lineup, they suggest that more research is directed on how to improve identification accuracy using a non-live lineup, regardless of the medium (photographs or video). As the current study was set in the United Kingdom, we followed U.K. guidelines where the preferred medium is the video lineup.

The materials used for the current study were developed and validated by Havard et al. (2012) and consisted of a short video of a staged theft committed by a young boy and the corresponding video lineups. The video footage shows a 9-year-old boy walking down a university corridor trying to open several locked office doors until he finds an unlocked one. He enters the office and after looking around, steals the occupant's purse, phone and laptop and then leaves the room. The video clip lasts 1 min 30 s, and the target is seen in

full face frontal and profile views throughout the film. We have used these materials due to the own-age bias in face recognition (Rhodes & Anastasi, 2012), as previous studies using adult targets and lineups may have disadvantaged child participants.

The identification lineups were presented in the form of a sequential video lineup. Both the TA and TP lineups contained nine targets presented one after the other for 15 s each. Each target was numbered in the top-left-hand corner of the video clip. Each video began with the target looking straight at the camera, then turning their head to the left before turning it to the right, then back to the front, with a short gap before the next target video began. These videos were recorded and approved at a VIPER (Video Identification Parades by Electronic Recording) suite and met specific criteria to ensure consistency in the recording of each video, which ensures that no video clip stands out from the lineup. Full details of the validation procedure can be found in Havard et al. (2012) where the materials were used with both child and adult participants.

### **Design**

The study employed a  $2 \times 2$  between-groups quasi-experimental design in which familiarity with the experimenter (familiar vs. unfamiliar) and lineup type (target-present vs. target-absent) were between-group factors. The dependent variables were the lineup identification accuracy and confidence rating. For TP lineups, there were four response types. Participants could respond with a correct identification of the target, a false identification (selection of a foil), incorrectly rejecting the lineup by saying that the target was 'not there' or indicating that they were 'not sure' by choosing the 'mystery boy'. For TA lineups, participants could respond by correctly selecting the 'not there' option, by indicating that they were not sure through their choice of the 'mystery boy' or by making an incorrect false identification by selecting the target replacement or any one of the foils. Confidence was



measured on a scale of 0–10 once participants had made their decision.

### **Procedure**

The study took place in two phases, and the procedure for familiar and unfamiliar participants was kept as similar as possible. In Phase 1, children viewed the short video of the staged theft in small groups of around 5–10 in a quiet area of their school or crèche. This format was consistent with that in similar studies (Havard & Memon, 2013; Havard et al., 2012).

The second phase took part 1–2 days later where each child carried out the identification task individually. They were asked if they remembered watching the film and were asked to describe what had happened but no recall data were recorded. All children were read the same instructions and then viewed either a TP or TA lineup. The exact instructions were as follows:

Hello, do you remember the video we watched the other day?

Can you tell me what happened in it? Do you remember what the person looked like?

In a moment, I am going to show you another video, and in this video, you are going to see nine different boys appear one at a time on the screen in front of you. If you see the boy from the video we watched the other day I would like you to draw a circle around that number on the sheet in front of you. But, he might not be one of the boys you see, and if you don't think he is any of them I would like you to circle 'not there'. If you can't remember or if you're not sure that's okay too, I would like you to circle the 'mystery boy' on the sheet in front of you. Does that make sense? Do you need me to say it again?

Would you like to watch the video now?

You can pause the video at any time, just let me know. And if you would like to see any of the boys again we can do that too.

Children watched the lineup twice in accordance with Police & Criminal Evidence Act 1984 (PACE) Codes of Practice (2017) and the Lord Advocates Guidelines (2007) and were told that they could go back and look at any of the individuals in the lineup again. The experimenter sat a distance from each child and made no eye contact and waited for them to fill in the answer sheet. There was no time limit for children to make their decision, and no feedback was provided. All the participants filled in their own response sheet. Once participants had made their choice, they were asked to indicate how sure they were about their decision and were informed that a low number on the scale meant that they were not very sure, a middle number meant that they were pretty sure and a high number meant that they were very sure about their answer. This stage of the study took around 10 minutes. After completing the lineup, the child was thanked for helping with the research and was debriefed.

### **Results**

Accuracy and confidence data were analysed separately for TP and TA lineups as identification decisions may be driven by different processes depending on the target's presence in the lineup (Pozzulo & Lindsay, 1999).

#### **Accuracy**

Table 1 shows the breakdown of identification responses by administrator familiarity for TP and TA lineups. On TP lineups, accuracy was slightly higher in the familiar group across the different types of response. A 2 (familiarity: familiar, unfamiliar)  $\times$  4 (response type: correct identification, false identification, not there, not sure) chi-square test of association showed that there was no significant relationship between experimenter familiarity and response type,  $\chi^2(3, N = 60) = 1.51, p = .681$ , on the TP lineups.

On TA lineups, accuracy was higher in the familiar group with more correct rejections (in

Table 1. Frequencies and percentages of identification response for target-present and target-absent lineups by administrator familiarity.

Response	Familiar		Unfamiliar	
	<i>n</i>	%	<i>n</i>	%
Target-present				
Correct ID	20 (0.3, 0.8)	67	17 (-0.3, -0.8)	57
False ID	4 (-0.2, -0.4)	13	5 (0.2, 0.4)	17
Not there	4 (-0.6, -1.0)	13	7 (0.6, 1.0)	23
Not sure (mystery boy)	2 (0.4, 0.6)	7	1 (-0.4, -0.6)	3
Target-absent				
False ID	11 (-0.4, -0.8)	36	14 (0.4, 0.8)	47
Not there	17 (1.1, 2.1)	57	7 (-1.1, -2.1)	23
Not sure (mystery boy)	2 (-1.2, -1.8)	7	9 (1.2, 1.8)	30

Note: Standardised and adjusted residuals in parentheses.

the form of ‘not there’ responses) and fewer false identifications and not sure responses. A 2 (familiarity: familiar, unfamiliar) × 3 (response type: false identification, not there, not sure) chi-square test of association revealed a significant association between experimenter familiarity and response type on TA trials,  $\chi^2(2, N = 60) = 5.60, p = .031$ , one-tailed, with a moderate effect size (Cramer’s  $V = .31$ ). Following guidance from the Cornell Statistical Consulting Unit (2020), inspection of the adjusted standardised residuals ( $\geq 1.96$ ) in Table 1 suggests that familiarity led to more ‘not there’ responses than expected by chance whereas the rate of false identifications did not appear to differ from what was expected by chance for either group. The results show that children were more likely to give a ‘not there’ response when the TA line-up was administered by a familiar adult than by an unfamiliar adult.

**Confidence**

Once participants had made their decision, they were asked to rate how confident they were in their decision on a scale from 0 to 10 (where 0 indicated that they were guessing, and 10 indicated that they were very confident that they were correct). On TP lineups, confidence ratings for correct and incorrect

responses were compared by classifying a response as correct when the target was correctly identified and incorrect when a participant selected another foil, the ‘not sure’ or a ‘not there’ response. On TA trials, ‘not there’ responses were classified as correct. In addition, ‘not sure’ responses were considered lineup rejections in line with the recommendation that police would classify ‘not sure’ responses as rejections (Pozzulo & Lindsay, 1997). Incorrect responses were classified as selection of the target replacement or the selection of a foil. Table 2 shows the mean confidence rating by administrator familiarity for TP and TA lineups.

On TP lineups, familiar and unfamiliar participants show a similar pattern, with higher confidence for correct than incorrect answers. The main effect of lineup accuracy was confirmed by a 2 (familiarity: familiar, unfamiliar) × 2 (lineup accuracy: correct vs incorrect) between-groups analysis of variance (ANOVA),  $F(1, 56) = 6.98, p = .011, \eta_p^2 = .11$ , with higher confidence when correct ( $M = 8.78, SD = 1.60, 95\% \text{ CI } [8.02, 9.51]$ , where CI = confidence interval) than when incorrect ( $M = 7.17, SD = 2.99, 95\% \text{ CI } [6.21, 8.12]$ ). The main effect of familiarity was not significant,  $F(1, 56) = 0.08, p = .780, \eta_p^2 = .001$ , and there was no significant interaction

Table 2. Mean confidence ratings and standard deviations on correct and incorrect responses for target-present and target-absent lineups by administrator familiarity.

Response	Familiar <i>M (SD)</i>	Unfamiliar <i>M (SD)</i>	Total <i>M (SD)</i>
Target-present			
Correct	9.00 (1.30)	8.53 (1.91)	8.78 (1.60)
Incorrect	7.10 (3.18)	7.23 (2.98)	7.17 (2.99)
Target-absent			
Correct	7.79 (2.68)	6.56 (2.85)	7.23 (2.79)
Incorrect	7.82 (2.04)	8.00 (1.71)	7.92 (1.82)

Note: Confidence ratings: 0–10. Standard deviations in parentheses.

between familiarity and accuracy,  $F(1, 56) = 0.25, p = .621, \eta_p^2 = .004$ .

On TA lineups, the familiar group showed no difference in their confidence ratings when correct or incorrect, and the unfamiliar group were slightly less confident when correct than incorrect. A 2 (familiarity: familiar, unfamiliar)  $\times$  2 (lineup accuracy: correct vs incorrect) between-groups ANOVA revealed no significant effects: main effect of familiarity,  $F(1, 56) = 0.67, p = .418, \eta_p^2 = .01$ ; main effect of response type,  $F(1, 56) = 1.31, p = .257, \eta_p^2 = .02$ ; interaction,  $F(1, 56) = 1.21, p = .276, \eta_p^2 = .02$ . Across both lineup types, confidence ratings were not impacted by experimenter familiarity.

## Discussion

The main aim of this study was to explore the role of familiarity with the lineup administrator on children's ability to correctly reject a target-absent (TA) lineup. The results showed that experimenter familiarity had an impact on lineup response in TA lineups only, with a tendency for more 'not there' and fewer 'not sure' responses when the children were familiar with the experimenter. Accuracy on target-present (TP) lineups was not influenced by experimenter familiarity, and there was no effect of familiarity on confidence on either TP or TA lineups. These results support our hypotheses and lend support to the role of social factors in children's high false

identification rate on TA lineups. This is consistent with previous work that found children's accuracy on TA lineups was impacted by social cues such as the administrator's clothing (Lowenstein et al., 2010) and supportive non-verbal cues (Rush et al., 2014) and similar work that has shown that young children are less susceptible to misleading questions when there is some degree of familiarity with the interviewer (Bjorklund et al., 2000; Quas & Schaaf, 2002). These findings have important implications for those working with child eyewitnesses and suggest that children may feel less pressure to choose someone from the lineup when there is some degree of familiarity with the lineup administrator. The level of familiarity may be a key factor here, and this warrants further investigation as studies have shown that children are less accurate when lineups were administered by their parents (Ricci et al., 1996) and that there is no effect of familiarity when they are interviewed a second time by the same person (Brubacher et al., 2019). The level of familiarity that the children had with the experimenter in the current study lies somewhere between these two extremes, and future research could examine the rate of false identifications when children have spent some time becoming familiar with the administrator before they make their lineup decision. Despite the improved accuracy on TA lineups with a familiar administrator, the improvement was more modest than some studies and should be interpreted cautiously.



The current findings do have important implications, however, as they do lend support to the role of social factors in children's eyewitness decisions and suggest a potential way in which the implicit pressure to choose someone could be reduced in a legal setting.

The current study also explored the effect on eyewitness accuracy when salient options were included in the lineup for both 'not there' and 'not sure' options (in the form of the 'mystery boy'). The inclusion of the 'not there' and the 'not sure' options in the same lineup did not have an overall effect on reducing false identifications but it seems to have helped the children who were unfamiliar with the experimenter to indicate uncertainty as they used the mystery boy option more often than the familiar children. These findings reinforce the potential benefits of including a 'mystery person' or wildcard option when presenting child eyewitnesses with an identification lineup and have direct implications for current identification procedures regardless of the medium of presentation (photo lineups as used in Australia, New Zealand and the United States or video lineups as preferred in the United Kingdom). However, it should be noted that the false identification rate for both groups of children was still relatively high, indicating that a lot of children still prefer to pick someone from the line-up rather than use the 'not sure/mystery boy' option. The impact of including both a 'not there' and a 'not sure' option within the lineup on reducing false identifications therefore needs to be fully explored and compared to a control condition with only a 'not there' option. The findings of the current study are limited by the lack of this control due to the constraints on participant numbers from our familiarity manipulation. Future research could explore the most effective way to present the 'not sure' option in the lineup for child witnesses and whether this should be in the form of the mystery person/wildcard and also consider the role of instructions to the child to try to respond as accurately as possible (Brewer et al., 2010).

Across a series of experiments combining 'not there' and 'not sure' options with adults, Lucas et al. (2020) found that the 'not sure' option was rarely selected and that the inclusion of an explicit 'not there' option within the lineup decreased false identifications and improved accuracy on TA lineups.

Familiarity with the experimenter did not have an effect on children's confidence in their decisions. Children were more confident when they were correct on TP lineups, and this is consistent with previous studies with both children and adults (Havard & Memon, 2013; Havard et al., 2012; D. S. Lindsay et al., 1998; Memon et al., 2003; Read, 1995) whereas there was no confidence-accuracy relationship for TA lineups. Overall, confidence ratings tended to be high, suggesting that the young children in our study are unable to reliably judge the accuracy of their answers and that meta-cognitive skills are still developing. Veenman et al. (2006) reported that metacognitive skills emerge at the ages of 8-10 years. While confidence may not be a reliable indicator of children's accuracy, it may be the case that there are other useful indicators of eyewitness accuracy. In a recent review of the eyewitness identification procedures, Wells et al. (2020) recommend that the entire lineup procedure should be video recorded to maintain an accurate account of the procedure and to ensure there is no administrator bias throughout. This recommendation is already noted in the identification guidelines in several Australian States (Cullen et al., 2021). We would argue that video recordings may also provide useful *non-verbal* indicators of children's recognition of the suspect. During the identification procedure in the current study, the researcher noticed that some children would respond non-verbally to the target when he appeared on screen, jumping up from their seat or nodding their head. We followed up the possibility that children could convey non-verbal cues that they had recognised the target in some exploratory pilot work. We video recorded a small group of children while they

watched either a TP or a TA lineup. There were clear differences in the non-verbal behaviour of this small group of children depending on which lineup they had viewed. The children who watched a TP lineup showed clear signs of disinterest once the target had appeared in the lineup – they got fed up, started to fidget and looked around the room. In contrast, the children who watched a TA lineup continued to pay close attention to all of the lineup members and showed no obvious non-verbal signs that they had recognised the target. Future work could explore whether non-verbal behaviour could be used as a reliable indicator of identification accuracy. Other implicit measures of recognition have been noted in a recent study by Winsor et al. (2021) where viewing behaviours such as the first face children chose to look at in the lineup and overall length of time spent interacting with the lineup predicted identification accuracy.

This study has some limitations. Firstly, we acknowledge that the study did not directly manipulate the level of familiarity with the experimenter. This was largely due to the practical difficulties of working with this age group and the time required to develop familiarity. In future, we would like to be in a position to randomly assign children from the same sample to the familiar or unfamiliar experimenter condition, but at the current time, this is not a practical request for schools as they continue to recover from lost teaching time over recent years. It is important to note that the children within each group were randomly assigned to the TP and TA conditions. Secondly, like a lot of other eyewitness research, there is only one target, and the results we have found could be limited to that target. Some researchers have included multiple lineups with the same participants; however, these are then constrained by time limits as they do not include an ecologically valid time delay between seeing the target and completing the lineup (e.g. Brewer et al., 2010). Participant numbers were restricted due to the familiarity manipulation, and so we do not have a control condition with only a ‘not

there’ option to fully explore the effects of including both ‘not there’ and ‘not sure’ options in the lineup. The familiarity manipulation also means that the age group was restricted to 5–7-year-olds, and while we found a beneficial effect of familiarity in this age group, we do not know whether these results will extend to older children. Finally, the lineup administrator in our study was not blind to which boy was the target and may potentially have provided unintentional cues to the participants – for example, a small smile when the participant recognised the target in the video (Charman & Quiroz, 2016; Zimmerman et al., 2017). However, this would apply to both familiar and unfamiliar children and is unlikely to have influenced the results, but this possibility strengthens the argument to video record the lineup procedure so that both the child’s and the administrator’s behaviour can be monitored for any potential non-verbal cues. During the second phase of the study, the lineups were administered individually, and due to time constraints this phase of the procedure was carried out 1–2 days after having watched the video. It is unknown whether there were any differences in accuracy between the children who completed the identification task one day after watching the video and those who completed it after two days as we did not record this information at the time of data collection. In future research, it would be interesting to investigate whether familiarity could mitigate against any decrease in accuracy due to the time delay between viewing the incident and making an identification decision as the evidence to date regarding the impact of a delay on identification accuracy with children is inconsistent (Clifford et al., 2012; Karageorge & Zajac, 2011).

Despite these shortcomings, the current study provides evidence towards a beneficial role of familiarity with the administrator on reducing false identifications in TA lineups and support for the role of social factors in children’s eyewitness decisions. The results suggest that the high error rate observed on

TA lineups in children is largely driven by a social pressure to choose someone, and young children are able to resist this social pressure and provide an accurate answer in some circumstances. Further research is required to correctly apply this in a legal setting. In addition, future work should explore different options of presenting a 'not sure' response for children and investigate children's non-verbal responses while watching the lineups as this offers a promising source of their recognition and more importantly their *lack* of recognition of a suspect.

### Acknowledgements

We would like to thank all the participants who took part in this study.

### Ethical standards

#### *Declaration of conflicts of interest*

Lesley Calderwood has declared no conflicts of interest

Carrie Ballantyne has declared no conflicts of interest

Kimberley Slee has declared no conflicts of interest

#### *Ethical approval*

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

#### *Informed consent*

Informed parental consent was obtained from all individual participants included in the study. Verbal assent was obtained from each participant on each day of testing.

### ORCID

Lesley Calderwood  <http://orcid.org/0000-0003-3608-4990>

Carrie Ballantyne  <http://orcid.org/0000-0002-8271-6787>

### Data availability statement

The authors would have liked to provide the data collected; however, permission for the data to be submitted to online repositories was not obtained from the parents and guardians of the participants at the time of consent.

### References

- Beal, C. R., Schmitt, K. L., & Dekle, D. J. (1995). Eyewitness identification of children. *Law and Human Behavior, 19*(2), 197–216. <https://doi.org/10.1007/BF01499325>
- Bjorklund, D. F., Cassel, W. S., Bjorklund, B. R., Brown, R. D., Park, C. L., Ernst, K., & Owen, F. A. (2000). Social demand characteristics in children's and adults' eyewitness memory and suggestibility: The effect of different interviewers on free recall and recognition. *Applied Cognitive Psychology, 14*(5), 421–433. [https://doi.org/10.1002/1099-0720\(200009\)14:5](https://doi.org/10.1002/1099-0720(200009)14:5)
- Brewer, N., & Day, K. (2005). The confidence-accuracy and decision latency-accuracy relationships in children's eyewitness identification. *Psychiatry, Psychology and Law, 12*(1), 119–128. <https://doi.org/10.1375/ppl.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*(2), 261–277. doi: <https://doi.org/10.1348/135532509X474822>
- Brewer, N., & Palmer, M. A. (2010). Eyewitness identification tests. *Legal and Criminological Psychology, 15*(1), 77–96. <https://doi.org/10.1348/135532509X414765>
- Brubacher, S. P., Poole, D. A., Dickinson, J. J., La Rooy, D., Szojka, Z. A., & Powell, M. B. (2019). Effects of interviewer familiarity and supportiveness on children's recall across repeated interviews. *Law and Human Behavior, 43*(6), 507–516. <https://doi.org/10.1037/lhb0000346>
- Ceci, S. J., Ross, D. F., & Toglia, M. P. (1987). Suggestibility of children's memory:

- Psycholegal implications. *Journal of Experimental Psychology: General*, 116(1), 38–49. <https://doi.org/10.1037/0096-3445.116.1.38>
- Charman, S. D., & Quiroz, V. (2016). Blind sequential lineup administration reduces both false identifications and confidence in those false identifications. *Law and Human Behavior*, 40(5), 477–487. <https://doi.org/10.1037/lhb0000197>
- Clifford, B. R., Havard, C., 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(1), 130–139. <https://doi.org/10.1002/acp.1804>
- Cornell Statistical Consulting Unit. (2020). *Using adjusted standardized residuals for interpreting contingency tables*. [https://cscu.cornell.edu/wp-content/uploads/95\\_conttbleresid.pdf](https://cscu.cornell.edu/wp-content/uploads/95_conttbleresid.pdf)
- Cullen, H. J., Adam, L., & van Golde, C. (2021). Evidence-based policing in Australia: An examination of the appropriateness and transparency of lineup identification and investigative interviewing practices. *International Journal of Police Science & Management*, 23(1), 85–98. <https://doi.org/10.1177/14613557211004618>
- Davies, G., Stevenson-Robb, Y., & Flin, R. (1988). Tales out of school: Children's memory for an unexpected event. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues, Vol. 1. Memory in everyday life* (pp. 122–127). John Wiley & Sons.
- Davis, S. L., & Bottoms, B. L. (2002). Effects of social support on children's eyewitness reports: A test of the underlying mechanism. *Law and Human Behavior*, 26(2), 185–215. <https://doi.org/10.1023/A:1014692009941>
- Dekle, D. J., Beal, C. R., Elliott, R., & Huneycutt, D. (1996). Children as witnesses: A comparison of lineup versus showup identification methods. *Applied Cognitive Psychology*, 10(1), 1–12. doi.org/%3C1::AID-ACP354%3E3.0.CO;2-Y [https://doi.org/10.1002/\(SICI\)1099-0720\(199602\)10:1](https://doi.org/10.1002/(SICI)1099-0720(199602)10:1)
- Dunlevy, J. R., & Cherryman, J. (2013). Target-absent eyewitness identification line-ups: Why do children like to choose. *Psychiatry, Psychology and Law*, 20(2), 284–293. <https://doi.org/10.1080/13218719.2012.671584>
- Evidence Act. (1995). *Australia*. Retrieved from [http://www6.austlii.edu.au/cgi-bin/viewdb/au/legis/cth/consol\\_act/ea199580/](http://www6.austlii.edu.au/cgi-bin/viewdb/au/legis/cth/consol_act/ea199580/)
- Evidence Act. (2006). *New Zealand*. Retrieved from <https://www.legislation.govt.nz/act/public/2006/0069/latest/DLM393463.html>
- Fitzgerald, R. J., & Price, H. L. (2015). Eyewitness identification across the life span: A meta-analysis of age differences. *Psychological bulletin*, 141(6), 1228–1265. <https://doi.org/10.1037/bul0000013>
- Fitzgerald, R. J., Price, H. L., & Valentine, T. (2018). Eyewitness identification: Live, photo, and video lineups. *Psychology, Public Policy, and Law*, 24(3), 307–325. <https://doi.org/10.1037/law0000164>
- Fitzgerald, R. J., Rubínová, E., & Juncu, S. (2021). Eyewitness identification around the world. In A. M. Smith, M. P. Toglia & J. M. Lampinen (Eds), *Methods, measures and theories in eyewitness identification tasks* (pp. 294–322). Routledge.
- Flvush, R., Hamond, N. R., Harsch, N., Singer, N., & Wolf, A. (1991). Content and consistency in young children's autobiographical recall. *Discourse Processes*, 14(3), 373–388. <https://doi.org/10.1080/01638539109544791>
- Havard, C. (2014). Are children less reliable at making visual identifications than adults? A review. *Psychology, Crime & Law*, 20(4), 372–388. <https://doi.org/10.1080/1068316X.2013.793334>
- Havard, C., & Memon, A. (2013). The Mystery Man can help to reduce false identification for child witnesses: Evidence from video lineups. *Applied Cognitive Psychology*, 27(1), 50–59. <https://doi.org/10.1002/acp.2870>
- Havard, C., Memon, A., Laybourn, P., & Cunningham, C. (2012). Own-age bias in video lineups: A comparison between children and adults. *Psychology, Crime & Law*, 18(10), 929–944. <https://doi.org/10.1080/1068316X.2011.598156>
- Jackson, S., & Crockenberg, S. (1998). A comparison of suggestibility in 4-year-old girls in response to parental or stranger misinformation. *Journal of Applied Developmental Psychology*, 19(4), 527–542. [https://doi.org/10.1016/S0193-3973\(99\)80054-9](https://doi.org/10.1016/S0193-3973(99)80054-9)
- 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 calibration. *Journal of Experimental Psychology: Learning Memory and Cognition*, 22, 1304–1316. doi/ <https://doi.org/10.1037/0278-7393.22.5.1304>
- Karageorge, A., & Zajac, R. (2011). Exploring the effects of age and delay on children's person identifications: Verbal descriptions,

- lineup performance, and the influence of wildcards. *British journal of Psychology (London, England : 1953)*, *102*(2), 161–183. <https://doi.org/10.1348/000712610X507902>
- Keast, A., Brewer, N., & Wells, G. L. (2007). Children's metacognitive judgments in an eyewitness identification task. *Journal of Experimental Child Psychology*, *97*(4), 286–314. <https://doi.org/10.1016/j.jecp.2007.01.007>
- Lindsay, R. C., 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*(4), 391–404. doi.apa.org/doi/ <https://doi.org/10.1023/A:1024807202926>
- Lindsay, D. S., Read, J. D., & Sharma, K. (1998). Accuracy and confidence in person identification: The relationship is strong when witnessing conditions vary widely. *Psychological Science*, *9*(3), 215–218. doi.org/10.1111%2F1467-9280.00041 <https://doi.org/10.1111/1467-9280.00041>
- Lord Advocates' Guidelines. (2007). *The conduct of visual identification procedures*. Retrieved from Lord Advocate's guidelines: Visual identification procedures | COPFS.
- 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*(1), 59–73. <https://doi.org/10.1002/jip.104>
- Lucas, C. A., Brewer, N., Michael, Z. E., & Foster, T. R. (2020). The effects of explicit "Not Present" and "Don't Know" response options on identification decisions in computer-administered lineups. *Applied Cognitive Psychology*, *34*(6), 1495–1509. <https://doi.org/10.1002/acp.3728>
- Memon, A., Bartlett, J., Rose, R., & Gray, C. D. (2003). The ageing eyewitness: The effects of face-age and delay upon younger and older observers. *Gerontology*, *58*, 338–345.
- Memon, A., Havard, C., Clifford, B., Gabbert, F., & Watt, M. (2011). A field evaluation of the VIPER system: A new technique for eliciting eyewitness identification evidence. *Psychology, Crime & Law*, *17*(8), 711–729. <https://doi.org/10.1080/10683160903524333>
- Memon, A., & Rose, R. (2002). Identification abilities of children: Does a verbal description hurt face recognition? *Psychology, Crime & Law*, *8*(3), 229–242. <https://doi.org/10.1080/10683160208401817>
- Mickes, L. (2015). Receiver operating characteristic analysis and confidence-accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory. *Journal of Applied Research in Memory and Cognition*, *4*(2), 93–102. <https://doi.org/10.1016/j.jarmac.2015.01.003>
- Olsson, N. (2000). A comparison of correlation, calibration, and diagnosticity as measures of the confidence-accuracy relationship in witness identification. *The Journal of Applied Psychology*, *85*(4), 504–511. doi/ <https://doi.org/10.1037/0021-9010.85.4.504>
- PACE (Police and Criminal Evidence Act 1984) Codes of Practice. (2017). Retrieved from Police and Criminal Evidence Act 1984 (PACE) codes of practice - GOV.UK ([www.gov.uk](http://www.gov.uk))
- Parker, J. F., & Carranza, L. E. (1989). Eyewitness testimony of children in target-present and target-absent lineups. *Law and Human Behavior*, *13*(2), 133–149. <https://doi.org/10.1007/BF01055920>
- Parker, J. F., & Ryan, V. (1993). An attempt to reduce guessing behavior in children's and adults' eyewitness identifications. *Law and Human Behavior*, *17*(1), 11–26. <https://doi.org/10.1007/BF01044534>
- Pica, E., Thompson, L. E., Pozzulo, J., Dempsey, J., & Pettalia, J. (2020). Variations of salient rejection options: Does one work best for adolescents and adults? *Journal of Police and Criminal Psychology*, *35*(4), 448–458. <https://doi.org/10.1007/s11896-018-9307-8>
- 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*(1), 25–34. <https://doi.org/10.1348/135532505X52626>
- Pozzulo, J. D., Dempsey, J., Bruer, K., & Sheahan, C. (2012). The culprit in target-absent lineups: Understanding young children's false positive responding. *Journal of Police and Criminal Psychology*, *27*(1), 55–62. <https://doi.org/10.1007/s11896-011-9089-8>
- Pozzulo, J. D., & Lindsay, R. C. L. (1997). Increasing correct identifications by children. *Expert Evidence*, *5*(4), 126–132. <https://doi.org/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*(5), 549–570.



- psycnet.apa.org/doi/ <https://doi.org/10.1037/0021-9010.84.2.167>
- Pozzulo, J. D., & Lindsay, R. C. L. (1999). Elimination lineups: An improved identification procedure for child eyewitnesses. *Journal of Applied Psychology, 84*(2), 167–176. psycnet.apa.org/doi/ <https://doi.org/10.1037/0021-9010.84.2.167>
- Quas, J. A., & Schaaf, J. M. (2002). Children's memories of experienced and nonexperienced events following repeated interviews. *Journal of Experimental Child Psychology, 83*(4), 304–338. [https://doi.org/10.1016/S0022-0965\(02\)00150-9](https://doi.org/10.1016/S0022-0965(02)00150-9)
- Raskin, D. C., & Yuille, J. C. (1989). Problems in evaluating interviews of children in sexual abuse cases. In S. J. Ceci, D. F. Ross, and M. P. Toglia (Eds.), *Children's eyewitness memory* (pp. 184207). New York: Springer Verlag.
- Read, J. D. (1995). The availability heuristic in person identification: The sometimes misleading consequences of enhanced contextual information. *Applied Cognitive Psychology, 9*(2), 91–121. <https://doi.org/10.1002/acp.2350090202>
- Ricci, C. M., Beal, C. R., & Dekle, D. J. (1996). The effect of parent versus unfamiliar interviewers on children's eyewitness memory and identification accuracy. *Law and Human Behavior, 20*(5), 483–500. <https://doi.org/10.1007/BF01499037>
- Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: a meta-analytic and theoretical review. *Psychological bulletin, 138*(1), 146–174. <https://doi.org/10.1037/a0025750>
- Rush, E. B., Quas, J. A., Yim, I. S., Nikolayev, M., Clark, S. E., & Larson, R. P. (2014). Stress, interviewer support, and children's eyewitness identification accuracy. *Child development, 85*(3), 1292–1305. [www.jstor.org/stable/24031926](http://www.jstor.org/stable/24031926)
- Sporer, S. L., Penrod, S. D., Read, J. D., & Cutler, B. L. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin, 118*(3), 315–327. doi/ <https://doi.org/10.1037/0033-2909.118.3.315>
- Thompson, L. E., Pozzulo, J., Pratt, K. J., Sheahan, C. L., Fraser, B. M., & Deweyert, C. E. (2020). Testing a new lineup procedure with children: The elimination with wildcard. *Journal of Applied Developmental Psychology, 69*, 101161. <https://doi.org/10.1016/j.appdev.2020.101161>
- Valentine, T., Darling, S., & Memon, A. (2007). Do strict rules and moving images increase the reliability of sequential identification procedures? *Applied Cognitive Psychology, 21*(7), 933–949. <https://doi.org/10.1002/acp.1306>
- Veenman, M. V., Van Hout-Wolters, B. H., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning, 1*(1), 3–14. <https://doi.org/10.1007/s11409-006-6893-0>
- Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. T. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law and Human Behavior, 44*(1), 3–36. <https://doi.org/10.1037/lhb0000359>
- Winsor, A. A., Flowe, H. D., Seale-Carlisle, T. M., Killeen, I. M., Hett, D., Jores, T., Ingham, M., Lee, B. P., Stevens, L. M., & Colloff, M. F. (2021). Child witness expressions of certainty are informative. *Journal of Experimental Psychology: General, 150*(11), 2387–2407. <https://doi.org/10.1037/xge0001049>
- Zajac, R., & Karageorge, A. (2009). The wildcard: A simple technique for improving children's lineup performance. *Applied Cognitive Psychology, 23*(3), 358–368. <https://doi.org/10.1002/acp.1511>
- Zimmerman, D. M., Chorn, J. A., Rhead, L. M., Evelo, A. J., & Kovera, M. B. (2017). Memory strength and lineup presentation moderate effects of administrator influence on mistaken identifications. *Journal of Experimental Psychology: Applied, 23*(4), 460–473. <https://doi.org/10.1037/xap0000147>