CrossMark

Developmental trends in lineup performance: Adolescents are more prone to innocent bystander misidentifications than children and adults

Nathalie Brackmann [1,2,3] · Melanie Sauerland [1] · Henry Otgaar [1,4]

Abstract

We tested developmental trends in eyewitness identification in biased and unbiased lineups. Our main interest was adolescent's lineup performance compared with children and adults. 7–10-year-olds, 11–13-year-olds, 14–16-year-olds, and adults ($N = 431$) watched a wallet-theft-video and subsequently identified the thief, victim, and witness from simultaneous target-present and target-absent six-person photo lineups. The thief-absent lineup included a bystander previously seen in thief proximity. Research on unconscious transference suggested a selection bias toward the bystander in adults and 11–13-year-olds, but not in younger children. Confirming our hypothesis, adolescents were more prone to bystander bias than all other age groups. This may be due to adolescents making more inferential errors than children, as predicted by fuzzy-trace theory and associative-activation theory, combined with lower inhibition control in adolescents compared with adults. We also replicated a clothing bias for all age groups and age-related performance differences in our unbiased lineups. Consistent with previous findings, participants were generally overconfident in their decisions, even though confidence was a better predictor of accuracy in older compared with younger participants. With this study, we show that adolescents have an increased tendency to misidentify an innocent bystander. Continued efforts are needed to disentangle how adolescents in comparison to other age groups perform in forensically relevant situations.

Keywords Identification performance · Unconscious transference · Child witnesses · Adolescent witnesses · Clothing bias · Confidence–accuracy relationship

Developmental differences in eyewitness memory oftentimes play a critical role in police investigations when conflicting statements are given by witnesses of different ages. In the realm of evaluating witnesses' credibility, young children's eyewitness memory is often deemed inferior to that of older witnesses (e.g., Bruck & Ceci, 1999). For eyewitness identification performance, we know that children, compared with adults, are either equally or less likely to correctly select a target from a target-present lineup, but are consistently less likely to correctly reject a lineup that does not include the culprit (see the two meta-analyses about age effects in

identification performance by Fitzgerald & Price, 2015, and Pozzulo & Lindsay, 1998). A more liberal response criterion has been suggested as the mechanism underlying this increased tendency to choose an innocent individual from target-absent lineups in younger children, possibly combined with a stronger tendency to guess, compared with older children and adults (e.g., Fitzgerald & Price, 2015).

In real-life police investigations, lineups generally include an individual the police suspects to be the culprit of the crime. The police may investigate the actual perpetrator (resulting in a target-present lineup) or mistakenly investigate an innocent individual (resulting in a target-absent lineup). Ideally, the witness would correctly reject a lineup that includes an innocent suspect. Age is one factor that can influence the rate at which innocent individuals are selected. Another impact factor concerns previous exposure to an innocent individual. Deffenbacher, Bornstein, and Penrod (2006) analyzed this effect in adults in a seminal meta-analysis. Selection rates of innocent individuals were found to be higher if they had been seen in a mugshot array that preceded the lineup or during the event itself as a bystander. This type of misidentification is

✉ Nathalie Brackmann
nathalie.brackmann@puk.zh.ch

1 Maastricht University, Maastricht, the Netherlands

2 University of Gothenburg, Gothenburg, Sweden

3 University Hospital of Psychiatry Zurich, Alleestrasse 61A, 8462 Rheinau, Switzerland

4 City, University of London, London, UK

called *unconscious transference,* or misidentification of a familiar bystander, and can be studied with two different experimental designs: either (1) all participants view the bystander in a target event, and subsequently the selection rate of the innocent bystander is compared with the selection rate of other (unfamiliar) foils, or (2) only half of the participants view a target event with the bystander present, and their selection rate of the innocent bystander is compared with the control participants who have not previously seen the innocent bystander. The mechanism underlying this effect may be that witnesses wrongly assume that the bystander and culprit are the same person. The mere presence of a familiar person in a simultaneous lineup may therefore increase the chance of his or her (erroneous) identification as the culprit.

Even though age effects are known to affect correct rejection rates, developmental research on innocent bystander misidentifications, which constitutes a special type of misidentification error, is lacking. One exception is a study by Ross et al. (2006), who compared bystander misidentifications in 5–12-year-olds. In this study, participants viewed one of two theft videos. The videos differed in only one scene, in which either a woman read to school children (control condition) or a man (the bystander) who had a similar appearance as the thief. Only the 11–12-year-olds and not the 5–10-year-olds who had seen the bystander in the video were more likely to erroneously select the bystander from a lineup with four foils compared with the age-matched control participants who had not previously seen the bystander. Therefore, a bystander bias was only apparent in the older children (11–12-year-olds).

In summary, bystander misidentifications have previously been studied in adults (e.g., Read, Tollestrup, Hammersley, McFadzen, & Christensen, 1990; Ross, Ceci, Dunning, & Toglia, 1994) and in younger children (Ross et al., 2006). However, to our knowledge, no study has thus far examined bystander misidentifications in adolescents. One important aim of the current study is to examine developmental trends in bystander misidentifications. The reason to focus on adolescents is a proposed likelihood of adolescents (people between ages 14 and 16) to make inferential errors that, in combination with underdeveloped executive functioning, makes them perform differently than younger children (younger than 10), older children (ages 11–13), and adults (as will be outlined below). The dearth of research on adolescent's eyewitness memory and identification performance not only pertains to the bystander misidentification effect but is a more general absence of research in the field of eyewitness memory that scholars have recently pointed out (Fitzgerald & Price, 2015; Jack, Leov, & Zajac, 2013) and have begun to challenge (McGuire, London, & Wright, 2015; Sauerland, Brackmann, & Otgaar, in press). For instance, Fitzgerald and Price could not compare adolescents with other age groups in their meta-analysis, because there were only three studies that included

this age group.[1] Also, other meta-analyses on eyewitness memory either focused exclusively on adults (Deffenbacher et al., 2006; Fitzgerald, Price, Oriet, & Charman, 2013; Sporer, Penrod, Read, & Cutler, 1995), or, if developmental trends were analyzed, did not include a comparison of adolescents with other age groups (Blank & Launay, 2014; Deffenbacher, Bornstein, Penrod, & McGorty, 2004; Köhnken, Milne, Memon, & Bull, 1999; Memon, Meissner, & Fraser, 2010; Pozzulo & Lindsay, 1998; Shapiro & Penrod, 1986; Steblay, Dysart, Fulero, & Lindsay, 2001). It is thus not only relevant to examine developmental specificities in bystander misidentifications, but also in eyewitness identification performance in general. Hence, our second aim is to compare developmental trends among children, adolescents, and adults in fair lineups that do not include an innocent bystander and bear best-practice lineup construction in mind (viz., Wells et al., 1998).

Examining adolescents is important because they form a unique age group whose brain and social maturation is under development (see also Jack et al., 2013). As outlined in Shulman et al. (2016), adolescents' decision-making is best described with a dual-system model in which reward sensitivity that promotes sensation seeking is at an interplay with cognitive control processes that aid self-regulation (see Fig. 1). The developmental difference between the two systems is greatest during late adolescence, accounting for difficulties in withholding prepotent reactions. This pattern is also supported by neuropsychological findings. More specifically, the prefrontal cortex (the dorsolateral regions in particular) that is associated with high-level executive control processes such as strategic self-monitoring continues to develop until roughly age 16 (Luciana, Conklin, Hooper, & Yarger, 2005). Furthermore, neural circuits that allow the inhibition of a prepotent response also undergo developmental changes throughout adolescence (Stevens, Kiehl, Pearlson, & Calhoun, 2007). Both processes are linked to metacognitive performance and the tendency to correct an error (Fernandez-Duque, Baird, & Posner, 2000). New research aimed to disentangle the factors behind the peak in adolescence in liberal decisions using an experimental risk-taking task in which ambiguity and uncertainty of choices were manipulated (van den Bos & Hertwig, 2017). That is, adolescents (15–16-year-olds), as compared with younger (8–14-year-olds) and older (17–22-year-olds) participants, seem to have a reduced ambiguity aversion and to search less for information that might reduce uncertainty.

---

[1] The three studies that were identified included adolescent comparison groups ages 13–15 (Havard, Memon, Clifford, & Gabbert, 2010), 12–14 (Pozzulo & Lindsay, 1997), and 10–14 years (Pozzulo & Warren, 2003). The studies differed in their designs, investigating sequential video versus static lineups (Havard et al., 2010) or simultaneous lineups (Pozzulo & Lindsay, 1997; Pozzulo & Warren, 2003) and showed no consistent age pattern. More specifically, there either was no age difference or a performance increase with age. None of these studies looked at bystander misidentifications.
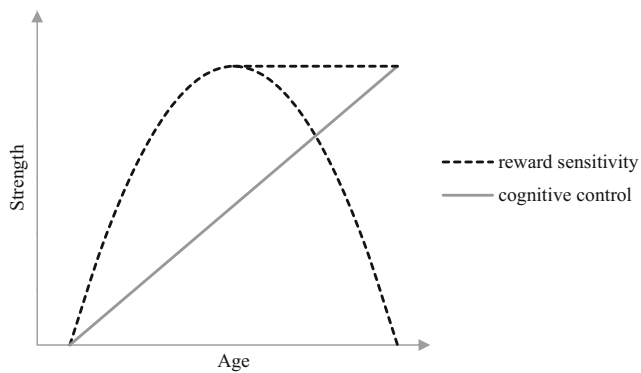
Fig. 1 Dual system model of decision-making from about age 10 to 25 as adapted from Shulman et al. (2016). Cognitive control may develop more or less steep or linear

These developmental changes may affect individuals' behavior in forensically relevant situations, including eyewitness identifications, leading to result patterns that differ from children and/or adults.

Theories that predict developmental differences in eyewitness memory such as lineup performance include fuzzy-trace theory (Brainerd, Reyna, & Ceci, 2008) and associative-activation theory (Howe, Wimmer, Gagnon, & Plumpton, 2009; Otgaar, Howe, Peters, Smeets, & Moritz, 2014). Both predict older children (e.g., 11–12-year-olds) and adults to be more likely than younger children (e.g., 7–8-year-olds) to make memory errors that are due to the retrieval of gist or relational information. This developmental memory effect, also known as *developmental reversal*, occurs because older children and adults have a more developed and dense knowledge base. Specifically, when children get older, they acquire more knowledge through experience and learning, which results in a more integrated and interrelated knowledge base. The result of this developmental improvement is that during encoding, associative activation is stronger and spreads faster to related concepts. During this spreading activation, incorrect associations can be made, and, hence, when such activation is stronger and more rapid, more incorrect associations arise. Furthermore, with increased age, the gist or meaning of events is more easily retrieved, leading to higher false-memory rates. Facial gist (resemblance in age, body build, etc.) is proposed to cause false-positive outcomes in identification parades (Brainerd et al., 2008). The misidentification of an innocent bystander might therefore qualify as an inferential memory error. Because children, compared with adults, make fewer automatic associations in their knowledge base, this may affect the unconscious transference effect. That is, the older participants are, the more inferential errors they make, because more and faster correct and incorrect associations will be made in their knowledge base (Howe et al., 2009). This should be reflected in an age-related increase in bystander misidentifications (see

also Ross et al., 2006). Such an age increase was found in the aforementioned study by Ross et al. (2006), in which their oldest age group (11–12-year-olds), but not the younger age groups (5–10-year-olds) demonstrated unconscious transference errors. The linear developmental trend of an age increase in inferential errors (see Brainerd, 2013, for an overview of diverging visual and nonvisual stimuli that elicited a developmental reversal) has recently also been extended to adolescents (McGuire et al., 2015; Quas et al., 2016), though it has not yet been tested in an eyewitness context such as lineup performance. To conclude, there seems to be a linear age increase in inferential errors, but adults might be better at second-guessing their decisions if additional information is available (which is reflected in improvements in cognitive control).

Adolescent's tendency toward liberal decision-making, in combination with a tendency to make inferential errors, may elevate adolescent's propensity to make bystander misidentification errors (even though the phenomenon is present in all ages). To test this in the current study, participants watched a wallet-theft video that showed the actual thief, the victim, an innocent bystander, and a witness. We hypothesized that when confronted with a bystander-present lineup that does not contain the actual target, adolescents would perform worse than children and adults. This is because of greater proneness of adolescents to inferential errors as predicted by fuzzy-trace theory and associative-activation theory and limited top-down control processes involved in inhibition control as derived from studies on brain maturation. In other words, adolescents should have the highest bystander misidentification rate compared with the other age groups due to a propensity to inferential errors (McGuire et al., 2015), riskier and more liberal decision-making processes (e.g., Gardner & Steinberg, 2005), and reduced inhibition control to withhold an answer (Stevens et al., 2007).

The tendency toward liberal decision-making and reduced inhibition control should also affect developmental trends in identification performance on a more global level. To investigate this, participants were also presented with a thief-present lineup and target-present or target-absent lineups concerning the victim and witness. These lineups did not include an innocent bystander, but tried to capture general developmental differences in the ability to make a correct lineup decision from unbiased lineups. For children and adults, we expected to replicate previous findings—namely, age increases in lineup performance for target/thief-present[2] (Fitzgerald & Price, 2015) and target/thief-

---

[2] Note that the prediction about target-present lineup performance is based on the recent meta-analysis by Fitzgerald and Price (2015). This deviates from the previous expectation of no age-related performance differences in target-present lineups, which were based on an earlier meta-analysis (Pozzulo & Lindsay, 1998).

absent lineups (Fitzgerald & Price, 2015; Pozzulo & Lindsay, 1998). For adolescents, again, research is lacking to make firm predictions. In line with the outlined liberal decision-making behavior, we expected an increased tendency to select a person from the lineups compared with adults, but similar to children.

For exploratory purposes, we also obtained confidence ratings for all lineup decisions. While confidence judgements are predictive of identification performance for adults' selections (but not rejections; see Sauerland & Sporer, 2009; Sporer et al., 1995; Weber & Brewer, 2004; Wixted, Read, & Lindsay, 2016), this does not seem to be the case for children (Brewer & Day, 2005; Keast, Brewer, & Wells, 2007). Specifically, children display overconfidence in their positive identification decisions. Again, knowledge is lacking about the confidence–accuracy relationship in adolescents (but see Brewer & Day, 2005, for a comparison between children and adolescents in target-present lineups). We can only speculate that the tendency toward impulse-driven answers in adolescents (Stevens et al., 2007) may result in overconfidence and hence a weaker confidence–accuracy relationship that is comparable to that of children. This would also be in line with the link between prefrontal cortex maturation (which is still under development in adolescents) and metacognitive performance (Fernandez-Duque et al., 2000), as metacognition may be seen as an indicator of calibration.

Method

Participants

Four-hundred-and-thirty-one participants, consisting of 98 7–10-year-olds (M = 8.53, SD = 0.87), 122 11–13-year-olds (M = 12.36, SD = 0.79), 100 14–16-year-olds (M = 14.90, SD = 0.67), and 111[3] adults (age range: 18–36 years, M = 22.45, SD = 2.84) participated in the experiment.[4] For child participants, consent of school principals and parents was obtained in addition to participants' consent. The study was approved by the standing ethical committee of the faculty. Adult participants (mainly undergraduate students) were granted study credit or were compensated with a 5€ gift voucher.

---

[3] Two adult participants ages 57 and 66 were excluded from the analyses because they were considered outliers (more than two standard deviations from the mean age of the adult sample).

Data has been made publicly available via the Open Science Framework at https://osf.io/58nhb/?view_only=f0010c7c2bae44a48bfeec5ac9290235.

[4] There is no clear consensus across studies about how adolescence is defined as age in years: While 11–13-year-olds are sometimes referred to as older children and other times as young adolescents, 14–16-year-olds are more uniformly labeled as adolescents (see Fig. 4 in van Duijvenvoorde, Peters, Braams, & Chrone, 2016). To account for these inconsistencies in the literature, we therefore chose four instead of three age groups to cover a broad range of nonadult ages.

Materials

Video

Participants were shown one of two versions of a 3-minute stimulus video depicting the nonviolent theft of a wallet. Both videos showed the same four actors: a thief, innocent bystander, victim, and witness. The videos differed in only two aspects: the appearance of the innocent bystander and the interaction of the innocent bystander with the victim. In Version A, the innocent bystander accidentally bumped into the victim before the theft took place. Furthermore, the bystander resembled the thief in appearance (same hair color, posture, color of clothes). In Version B, the victim passed the innocent bystander without making physical contact, and the bystander's clothes differed from those of the thief.[5] All actors were seen close up and from a distance. It was ensured that the bystander and thief were seen for the same amount of time (~25 seconds from close up and ~20 seconds from a distance). The comparable exposure duration and the moderate similarity in appearance was established to control for estimator variables that may play a role in the unconscious transference effect (Read et al., 1990; Ross et al., 1994; Ross et al., 2006).

Lineups and lineup construction Six lineups (three target present, three target absent, one of which was bystander present) were constructed for the male thief/bystander, the female victim, and a male witness (aged 22 to 26). Target presence (target-present vs. target-absent/bystander-present lineup presentation) was fully counterbalanced between participants.

Each lineup consisted of six 8.4 × 7.2-cm shoulder-up photos labeled A to F that were arranged in two rows of three pictures (a simultaneous lineup). The target position for the thief lineup was B, for the victim lineup C, and E for the witness lineup. The target position was chosen randomly during lineup construction. The fillers were the same within lineups. Effective sizes for the lineups, determined as Tredoux's Es, were high with a range of 4.1 to 5.6 (Tredoux, 1998, 1999), as established in a pilot study in which each lineup accompanied by a description of the target was presented to 19 to 38 individuals.

Design and procedure

The study used a 4 (age: 7–10-year-olds, 11–13-year-olds, 14–16-year-olds, and adults) × 3 (lineup: thief/bystander, victim, and witness) × 2 (presence: target-present vs. target-absent/bystander-present) mixed-factorial design. Age and

---

[5] Based on theoretical expectations derived from fuzzy-trace theory and associative-activation theory, we expected more bystander misidentifications for Version A than from B. However, as no effect of video version was found, we will not discuss this factor further.

presence served as between-subject variables, while lineup was a within-subject variable. Identification accuracy (accurate vs. inaccurate) served as the dependent variable and was defined as the proportion of correct decisions across all lineup decisions.

Participants were tested individually in quiet rooms, and test sessions lasted approximately 15 minutes. Participants viewed the stimulus video, completed a short filler task (2 minutes of *finding the differences between two pictures*), and then made the three identification decisions. Participants were first asked by the experimenter to identify the person who stole the wallet (thief/bystander lineup) and were then presented with the victim and the witness lineup (always in this order). They could either make a selection, state that the target was not in the lineup (lineup rejection), or indicate that they did not know. Prior to the presentation of the lineups, participants were informed that the targets may or may not be present in the lineup. Each lineup decision was followed by the question of whether any other, nonidentified lineup member had been present in the video. If so, participants were asked in which scene they had seen this person. Following each identification decision, participants indicated their postdecision confidence on an 11-point scale ranging from 0% to 100%. The scale was accompanied by smileys (ranging from a sad face corresponding to 0% confidence to a happy face corresponding to 100% confidence) to facilitate children's choices. No postdecision confidence ratings were obtained for "don't know" responses. After the test session, participants were fully debriefed.

Results

The frequencies of participants' identification responses, differentiated by age, for the thief lineups are depicted in Table 1; the referring data for victim and witness lineups can be found in Table 2. Across lineups and age groups, 42% of participants made a correct decision (correct identifications and correct rejections). We performed binomial logistic regressions (simultaneous entry) to establish the effect of age on the likelihood of different lineup outcomes (correct identification, foil identification, correct and incorrect rejections, don't-know response) for each target. When a significant age effect was found, post hoc comparisons were performed. Only significant comparisons are reported. The referring statistics will not be reported in the text, but can be found in Tables 3–5.

Two ways of handling don't-know responses were possible: treating them as missing values or coding them as lineup rejections (see Sauerland et al., 2016, for a similar approach). We conducted both analyses, and this resulted in analogous outcomes. The results reported in the main text refer to analyses treating don't-know responses as missing values. Participants rarely gave a don't-know response, and age did not have an effect on don't-know response rates for any of the lineups, Wald χ^2s(3) $\leq$ 2.39, ps $\geq$ .495.

Developmental differences in innocent bystander misidentifications

Bystander-present/thief-absent lineup For misidentifications of the bystander as the perpetrator in the bystander-present/thief-absent lineup, χ^2(3, $N = 219$) = 19.61, $p < .001$; Nagelkerke's R^2 = 13.2%, correct classification rate of 78.1%, post hoc comparisons showed that adolescents (14–16-year-olds) misidentified the bystander more often than all other age groups (see bottom part of Tables 1 and 3). Across ages, in response to the follow-up question if any other lineup member could be recognized from the video, 22.9% ($n = 50$) of the participants indicated to have seen the innocent bystander in the video and therefore correctly identified him as the bystander rather than the thief. There were no age differences in follow-up identifications of the bystander, Wald χ^2(3) = 4.56, $p = .207$.

Age did affect foil selections[6] from the bystander-present/thief-absent lineup, χ^2(3) = 12.53, $p = .006$ (Nagelkerke's R^2 = 7.4%, correct classification rate of 62.1%). Post hoc comparisons showed that young children (7–10-year-olds) were more likely to select a foil than were all other age groups. Age did not affect correct rejections from the bystander-present/thief-absent lineup, Wald χ^2(3) = 6.04, $p = .110$.

Developmental differences in identification performance in general

Thief-present/bystander-absent lineup For correct identifications from the thief-present/bystander-absent lineups, the logistic regression model was statistically significant, χ^2(3, $N = 213$) = 10.00, $p = .019$. The model explained 6.3% (Nagelkerke's R^2) of the variance and correctly classified 63.8% of the cases. Post hoc comparisons showed that adults were less likely to correctly identify the thief than were 11–13-year-olds and 14–16-year-olds (see top part of Tables 1 and 3). Age affected neither foil selections from the thief-present/bystander-absent lineup, Wald χ^2(3) = 2.90, $p = .407$, nor false rejections, Wald χ^2s(3) = 7.11, $p = .069$.

Victim lineup For correct identifications from the victim-present lineup, χ^2(3, $N = 211$) = 13.32, $p = .004$; Nagelkerke's R^2 = 8.4%, correct classification rate of 63.0%, the two younger age groups differed significantly from both

---

[6] Note that the identification of the bystander was not considered a foil selection.

Table 1 Thief-present/bystander-absent and thief-absent/bystander-present lineup: Percentages of identification outcomes as a function of age

	Age group	n	Thief identifications (correct decision)	Foil selections	Lineup rejections	Don't know responses
Thief-present/bystander-absent lineup	7–10	50	60.0	30.0	10.0	0.0
	11–13	57	68.4[a]	26.3	5.3	0.0
	14–16	50	78.0[b]	16.0	6.0	0.0
	Adults	56	50.0[ab]	26.8	19.6	3.6
			Bystander misidentifications	Foil selections	Lineup rejections (correct decision)	Don't know responses
Bystander-present/thief-absent lineup	7–10	48	8.3[c]	64.6[fgh]	22.9	4.2
	11–13	65	16.9[d]	41.5[f]	41.5	0.0
	14–16	50	44.0[cde]	30.0[g]	24.0	2.0
	Adults	57	19.6[e]	42.9[h]	28.6	8.9

Note. Rows sharing the same superscript letter differ significantly ($p < .05$; see Table 3 for inferential statistics)

older age groups, with younger age groups being less likely to correctly identify the victim (see Tables 2 and 4).

The same age pattern emerged for foil selections from the victim-present lineup, with higher odds of incorrectly selecting a foil from the lineup in the two younger age groups compared with older participants, $\chi^2(3) = 20.35$, $p < .001$ (Nagelkerke's $R^2 = 13.5\%$, correct classification rate of 74.4%). Similarly, young children (7–10-year-olds) were more likely to select a foil[7] from the victim-absent lineup than the 14–16-year-olds and adults were, $\chi^2(3, N = 222) = 13.76$, $p = .003$ (Nagelkerke's $R^2 = 8.1\%$, correct classification rate of 60.8%).

Age did not affect false rejections from the victim-present lineup, Wald $\chi^2(3) = 0.28$, $p = .963$, but did have an effect on correct rejections from the victim-absent lineup, $\chi^2(3) = 13.89$, $p = .003$ (Nagelkerke's $R^2 = 8.2\%$, correct classification rate 60.4%). Young children (7–10-year-olds) were less likely to reject the target-absent lineup, compared with all other age groups.

Witness lineup Correct identifications from the witness-present lineup, $\chi^2(3, N = 217) = 26.14$, $p < .001$ (Nagelkerke's $R^2 = 15.6\%$, correct classification rate of 66.8%) occurred significantly less often in young children (7–10-year-olds) than in all other age groups. Additionally, adults made significantly more correct identifications than 14–16-olds (see Tables 2 and 5).

Foil selections from the witness-present lineup, $\chi^2(3) = 27.83$, $p < .001$ (Nagelkerke's $R^2 = 16.1\%$, correct classification rate of 65.9%), occurred significantly more often in young children (7–10-year-olds) than in any other age group. No effect was found for foil selections from the witness-absent lineup, Wald $\chi^2(3) = 2.13$, $p = .545$. Age did not have a

significant effect on false rejections, Wald $\chi^2(3) = 5.89$, $p = .117$, or correct rejections of the witness lineups, Wald $\chi^2(3) = 2.99$, $p = .393$.

Comparing the witness lineup results to the victim lineup results revealed relatively high rates of foil selections across age groups (see Table 2) for the witness lineup, also compared with typical findings in the identification literature (e.g., Fitzgerald & Price, 2015). Another inspection of the lineup decisions revealed a clothing bias toward one of the foils (Position B) in the witness lineups. We will revisit this point in the Discussion.

Confidence–accuracy relationship

In order to generate choosers' calibration curves with a sufficient n per category, we collapsed the 11 confidence categories into four categories and report the weighted averages (0%–40%, 50%–60%, 70%–80%, and 90%–100%; see, e.g., Juslin, Olsson, & Winman, 1996, for a similar approach). They can be found in Fig. 2. The calibration curve allows the visual inspection of the relationship between participants' confidence judgments (the mean confidence across participants plotted on the x-axis) and accuracy (the proportion correct for each of the collapsed confidence categories plotted on the y-axis). For a detailed description of the measures used to assess participants' calibrations see Brewer and Wells (2006). A calibration curve above the ideal curve reflects underconfidence, while a calibration curve under the ideal reflects overconfidence. For example, the ideal outcome in the confidence category of 70% would be that across participants, 70% would give a correct response. In reality, however, only 50% might be accurate, reflecting overconfidence in the accuracy of lineup decisions.

The current data reveal little slope for the youngest age group across all three lineups. The same can be said for

---

[7] Note that we did not designate an innocent suspect in our target-absent lineups.

Table 2 Target-present and target-absent victim and witness lineups: Percentages of identification outcomes as a function of age

	Age group	n	Target present				n	Target absent		
			Target identifications (correct decision)	Foil selections	Lineup rejections	Don't know responses		Foil selections	Lineup rejections (correct decision)	Don't know responses
Victim lineup	7–10	48	50.0[ab]	39.6[ef]	10.4	0.0	50	76.0[ij]	22.0[klm]	2.0
	11–13	59	52.5[cd]	37.3[gh]	10.2	0.0	63	60.3	39.7[k]	0.0
	14–16	50	72.0[ac]	16.0[eg]	12.0	0.0	50	48.0[i]	50.0[l]	2.0
	Adults	54	77.8[bd]	9.3[fh]	13.0	0.0	59	44.1[j]	54.2[m]	1.7
Witness lineup	7–10	49	10.2[nop]	77.6[rst]	12.2	0.0	49	85.7	12.2	2.0
	11–13	60	38.3[n]	45.0[r]	13.3	3.3	62	80.6	17.7	1.6
	14–16	50	36.0[oq]	36.0[s]	28.0	0.0	50	82.0	18.0	0.0
	Adults	58	55.2[pq]	31.0[t]	13.8	0.0	55	74.5	25.5	0.0

Note. Rows sharing the same superscript letter differ significantly ($p < .05$; see Tables 4 and 5 for inferential statistics)

the witness lineup, regardless of age. This reflects a weak relationship between confidence and accuracy for young witnesses and the witness lineup. The remaining curves display a positive linear confidence–accuracy relationship. Of note, all obtained calibration curves lay under the ideal calibration curves, indicating that, overall, participants were overconfident.

The calibration statistics can be found in Table 6, and results are analogous to the above discussed inspection of the calibration curves. The C statistic can vary between 0 and 1 (with zero reflecting perfect calibration). Calibration was poor

for the witness lineup across all ages. This is not surprising bearing in mind that a clothing bias was found in the witness lineup. Thus, the calibration data adds as evidence to this unexpected finding. In the absence of clothing bias (i.e., ignoring the witness lineup), adolescents were relatively well calibrated (C $\leq$ .052), but in the presence of clothing bias calibration was poor (C = .199). The youngest age group displayed poorest calibration across lineups (C $\geq$ .131). The overconfidence/underconfidence statistic (O/U), which can vary between −1 (underconfidence) to +1 (overconfidence), indicated overconfidence for all slopes. No clear pattern of

Table 3 Inferential statistics of the post hoc comparisons between age groups for the thief lineup

		B	SE	Wald	df	p	Odds ratio	95% CI for odds ratio	
								Lower	Upper
Thief-present/bystander-absent lineup									
Thief identifications (correct decision)	Age			9.49	3	.023			
	11–13 vs. adults	.77	.39	3.92	1	.048	2.17	1.01	4.66
	14–16 vs. adults	1.27	.43	8.52	1	.004	3.55	1.52	8.29
	Constant	.61	.15	16.95	1	.000	1.84		
Bystander-present/thief-absent lineup									
Bystander identifications	Age			18.25	3	.000			
	7–10 vs. 14–16	−2.16	.60	13.14	1	.000	.12	.04	.37
	11–13 vs. 14–16	−1.35	.44	9.56	1	.002	.26	.11	.61
	14–16 vs. adults	−1.17	.44	7.02	1	.008	.31	.13	.74
	Constant	−1.41	.19	55.16	1	.000	.24		
Foil selections	Age			11.82	3	.008			
	7–10 vs. 11–13	−.94	.39	5.75	1	.016	.39	.18	.84
	7–10 vs. 14–16	−1.45	.43	11.25	1	.001	.24	.10	.55
	7–10 vs. adults	−.89	.41	4.81	1	.028	.41	.19	.91
	Constant	−.22	.14	2.38	1	.123	.80		

Note. Only significant post hoc comparisons are reported

Table 4 Inferential statistics of the post hoc comparisons between age groups for the victim lineup

		B	SE	Wald	df	p	Odds ratio	95% CI for odds ratio	
								Lower	Upper
Victim-present lineup									
Victim identifications	Age			12.61	3	.006			
	7–10 vs. 14–16	.94	.43	4.89	1	.027	2.57	1.11	5.94
	7–10 vs. adults	1.25	.44	8.24	1	.004	3.50	1.49	8.23
	11–13 vs. 14–16	.84	.41	4.25	1	.039	2.32	1.04	5.18
	11–13 vs. adults	1.15	.42	7.57	1	.006	3.16	1.39	7.18
	Constant	.58	.15	14.78	1	.000	1.78		
Foil selections	Age			17.13	3	.001			
	7–10 vs. 14–16	−1.24	.49	6.47	1	.011	.29	.11	.75
	7–10 vs. adults	−1.86	.56	11.24	1	.001	.16	.05	.46
	11–13 vs. 14–16	−1.14	.47	5.86	1	.016	.32	.13	.81
	11–13 vs. adults	−1.76	.54	10.61	1	.001	.17	.06	.50
	Constant	−1.22	.18	45.10	1	.000	.30		
Victim-absent lineup									
Foil selections	Age			12.69	3	.005			
	7–10 vs. 14–16	−1.23	.44	8.01	1	.005	.29	.12	.69
	7–10 vs. adults	−1.39	.42	10.85	1	.001	.25	.11	.57
	Constant	.31	.14	4.83	1	.028	1.37		
Rejections (correct choice)	Age			12.57	3	.006			
	7–10 vs. 11–13	.85	.43	3.92	1	.048	2.33	1.01	5.39
	7–10 vs. 14–16	1.27	.44	8.15	1	.004	3.55	1.49	8.45
	7–10 vs. adults	1.44	.43	11.15	1	.001	4.20	1.81	9.76
	Constant	−.38	.14	6.93	1	.008	.69		

Note. Only significant post hoc comparisons are reported

Table 5 Inferential statistics of the post hoc comparisons between age groups for the witness lineup

		B	SE	Wald	df	p	Odds Ratio	95% CI for odds ratio	
								Lower	Upper
Witness-present lineup									
Witness identifications	Age			19.75	3	.000			
	7–10 vs. 11–13	1.70	.54	9.85	1	.002	5.47	1.89	15.81
	7–10 vs. 14–16	1.60	.56	8.26	1	.004	4.95	1.66	14.73
	7–10 vs. adults	2.38	.54	19.41	1	.000	10.83	3.75	31.26
	14–16 vs. adults	.78	.40	3.92	1	.048	2.19	1.01	4.75
	Constant	−.75	.17	20.25	1	.000	.47		
Foil selections	Age			23.75	3	.000			
	7–10 vs. 11–13	−1.44	.43	11.24	1	.001	.24	.10	.55
	7–10 vs. 14–16	−1.82	.45	16.15	1	.000	.16	.07	.40
	7–10 vs. adults	−2.04	.45	21.00	1	.000	.13	.05	.31
	Constant	−.08	.15	.32	1	.572	.92		

Note. Only significant post hoc comparisons are reported. No significant post hoc comparisons emerged for the witness-absent lineup
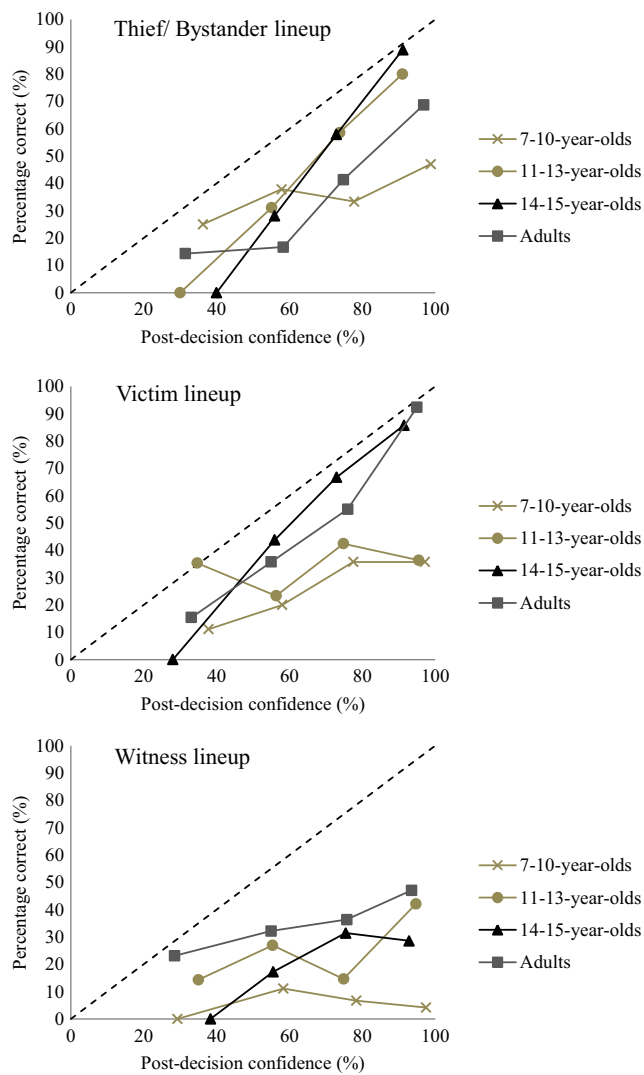
Fig. 2 Chooser's calibration curves for the thief/bystander, victim, and witness lineups. Dotted lines denote ideal calibration

Table 6 Confidence calibration (C), over/underconfidence (O/U), normalized resolution index (NRI), and point-biserial confidence–accuracy correlation (r) for three lineups differentiated by age group

Age group	C	O/U	NRI	r
Thief/bystander lineup				
7–10	.131	.314	.083	.136[**]
11–13	.056	.203	.234	.425[**]
14–16	.052	.200	.238	.487[**]
Adults	.115	.313	.197	.362[**]
Victim lineup				
7–10	.230	.463	.052	.179[**]
11–13	.129	.082	.297	.078[**]
14–16	.027	.106	.232	.425[**]
Adults	.047	.136	.470	.609[**]
Witness lineup				
7–10	.497	.092	.661	<.001[**]
11–13	.239	.450	.111	.129[**]
14–16	.199	.432	.067	.184[**]
Adults	.124	.034	.308	.119[**]

[**]$p \leq .01$

to the witness lineup (which displayed the aforementioned clothing bias), for which discriminative power was weak.

Discussion

The present study examined developmental trends in lineup performance. Child, adolescent, and adult participants consecutively viewed a thief, a victim, and a witness lineup after having watched a wallet-theft video. Half of these participants were confronted with a biased lineup that contained a familiar, but innocent, bystander. All other lineups were intended to be fair. For the biased bystander lineup, we expected adolescents to have the highest bystander misidentification rate compared to all other ages. This prediction was based on brain imaging studies indicating neurological changes involving inhibition control during adolescence (Stevens et al., 2007), a linear age increase in inferential errors (as predicted by fuzzy-trace theory and associative-activation theory; Brainerd et al., 2008; Howe et al., 2009; McGuire et al., 2015; Otgaar et al., 2014), and a tendency toward liberal decision-making in adolescents (Gardner & Steinberg, 2005). Confirming this hypothesis, adolescents were more likely than the other age groups to erroneously misidentify an innocent bystander as the thief. Furthermore, for the lineups that did not include an innocent bystander (i.e., victim and witness lineups, and thief-present lineup), we hypothesized age increases in accuracy for target-present and target-absent lineups

overconfidence/underconfidence as a function of age or lineup type was apparent.

The normalized resolution index (NRI) informs us on how well participants' confidence judgments discriminate accurate from inaccurate decisions. It can range between 0 and 1 (perfect discrimination). All NRIs were moderate to large (with the cutoff for a small effect being .010, for a moderate .059, and for a large .138; Brewer & Wells, 2006). Thus, even though participants were in general overconfident, confidence was still an indicator for accuracy. More specifically, we found large discriminability effects for adults across lineups and moderate effects for the youngest age group (with the exception of the witness lineup for which the effect was large). For older children (11–13 years) and adolescents, the NRI statistic revealed a strong capability of confidence to discriminate between accurate and inaccurate selections, except when it came

(Fitzgerald & Price, 2015). As predicted, older participants outperformed younger participants in the victim lineups and witness-present lineup, while, unexpectedly, this pattern was reversed in the thief-present lineup. Another unexpected finding was the high rates of foil selections across all age groups in the witness lineup, possibly reflecting a clothing bias. We will now discuss the details and relevance of these results in the following paragraphs.

One of our research aims was focused on innocent bystander identifications. The most interesting finding was that 14–16-year-old adolescents were more prone to the unconscious transference error than were all other age groups. This was reflected in higher rates of innocent bystander selections from the thief-absent but bystander-present lineup as compared with children and adults. This finding expands prior work that typically tested either children or adults, but not both, and never adolescents (Read et al., 1990; Ross et al., 1994; Ross et al., 2006). This pattern of results may be due to an interplay of two mechanisms: (1) developmental changes in brain maturity that are associated with an underdeveloped knowledge base (as postulated by fuzzy-trace theory and associative-activation theory; Brainerd et al., 2008; Howe et al., 2009) and reduced executive control processes (Luciana et al., 2005), and (2) ongoing developmental changes in social processes that lead to risky or liberal and impulsive lineup decisions (Fitzgerald & Price, 2015; Gardner & Steinberg, 2005).

The increased bystander bias of adolescents compared with the younger age groups partially supports the idea of developmental reversal (Brainerd et al., 2008). Developmental reversal has previously been found using paradigms eliciting spontaneous inferential errors (e.g., McGuire et al., 2015). We demonstrate an applied example of an inferential error that is particularly prevalent in adolescents and that confirms both that developmental trends in memory are flexible and that false memories are sometimes even more likely to arise with age (Otgaar, Howe, Brackmann, & Smeets, 2016). Unlike several studies that demonstrated a linear trend throughout adulthood in inferential errors, and what is predicted by fuzzy-trace theory and associative-activation theory (Brainerd et al., 2008), the bystander misidentification error in the current study again reduced from adolescents to adults. This supports the argument that an interplay of brain maturity processes and inhibition control processes may lead to elevated errors in adolescents compared with children, but reduced rates of errors in adolescents relative to adults. That is, being confronted with the picture of a previously seen person may have led to automatic associations within the knowledge base and therefore may have aided in recognition of that person and other related persons. While an age increase in automaticity of such inferences sheds light on superior face recognition in general, it also explains the

consequence of incorrect associations—in this case, the bystander misidentification. Strategic self-monitoring that is associated with maturation of the prefrontal cortex may have protected adults from this memory error while these top-down processes may not yet be as developed in adolescents (Luciana et al., 2005).[8] To our knowledge, no previous study has considered bystander misidentification rates in adolescents, and our study for the first time reports elevated bystander misidentifications in adolescents.

We also considered lineups that did not include an innocent bystander. Here, we found an age increase in accuracy for the fair victim lineup, as expressed in higher correct victim identification rates in both older age groups (adults and 14–16-year-olds) as opposed to both younger age groups (7–10-year-olds and 11–13-year-olds). This finding is in line with developmental patterns in foil selections—namely, that there are more foil selections in younger compared with older participants. Furthermore, the youngest age group was less likely than all other age groups to correctly reject the victim-absent lineup, which was again in line with increased foil selection rates. These results corroborate the meta-analytical findings of Fitzgerald and Price (2015). Another age pattern was found for the unbiased thief-present lineup. Surprisingly, adults made significantly fewer correct thief identifications than did 11–13- and 14–16-year-olds. Drawing from a substantial body of research showing adults' superiority in target-presence performance (see meta-analyses by Fitzgerald & Price, 2015, and Pozzulo & Lindsay, 1998), we believe that this finding may reflect a false positive outcome (Type I error) rather than a true age decline.

An unanticipated finding was that all age groups frequently selected a foil from the witness lineup, while the correct decision to select the witness or to reject the lineup was rarely chosen. Although we aimed for constructing fair lineups by piloting them prior to testing, a clothing bias toward one of the foils in the witness lineup was discovered post hoc. This foil, unfortunately, not only fit the general description of the witness (as intended) but also wore a red sweater similar to the one worn by the witness in the video. Our findings therefore, unintendedly, replicate seminal work on clothing biases (e.g., Lindsay, Wallbridge, & Drennan, 1987) that has led to adjustments in the best-practice recommendations on lineup constructions (Wells et al., 1998). Interestingly, notwithstanding that the clothing bias negatively affected all age groups, the youngest age group performed the worst and adults the best in the witness-present lineup. This age pattern is similar to the unbiased victim lineup described above and in accordance

---

[8] In line with this argumentation, one might assume that adults would have excelled in correct bystander identification in response to the follow-up question if any other, nonidentified lineup member had been present in the video. Our data do not support this hypothesis, though we might have had insufficient power to have detected such an age effect (only 50 participants made a post hoc identification of the bystander).

with the predictions derived from Fitzgerald and Price (2015). Our results therefore not only speak to age effects in unbiased lineups and lineups including an innocent bystander, but, unintentionally, also to age effects in lineups that display a clothing bias.

In summary, for unbiased lineups, adults generally outperformed the younger age groups on correct target identifications and lineup rejections (with the exceptions of the thief-present lineup in which adults actually were the least likely to correctly select the thief). The youngest age group was most prone to the influence of biases in the sense that they displayed an increased tendency to select foils if a clothing or bystander bias was present. However, adolescents, compared with all older and younger age groups, were most likely to misidentify a familiar bystander. Thus, even though the youngest age group was also more likely to select a foil, the foil was more randomly chosen by the 7–10-year-olds and not biased toward the innocent bystander (as it was in adolescents). Apparently, the liberal decision-making only influences lineup performance if wrong inferences about the bystander's involvement in the crime are drawn.

Finally, exploratory analyses investigated the confidence–accuracy relationship of choosers across the four age groups. Replicating earlier findings, all age groups displayed considerable overconfidence (Brewer & Day, 2005; Weber & Brewer, 2004). Also in line with previous findings, confidence and accuracy were largely unrelated for young children below the age of 11 (Brewer & Day, 2005). From the age of 11 onwards, a positive linear confidence accuracy was apparent for the thief/bystander and victim lineups, but not the witness lineup. This latter finding can most likely be ascribed to the aforementioned clothing bias. Adolescents' calibration curves approached the ideal calibration curves most closely, in comparison to all other age groups. This indicates that adolescents, though overconfident, were able to give a fair estimate of their likelihood to have made a correct decision. Even though overconfidence is in line with our prediction based on reduced inhibition control in adolescents (Stevens et al., 2007), this does not weaken the confidence–accuracy calibration overall. When evaluating adolescents' lineup performance, one may therefore cautiously draw conclusions that there is a higher chance of correct target selection when confidence is high. Because there is, to our knowledge, no other study that has investigated the confidence–accuracy relationship in adolescents in target-present and target-absent lineups (but see Brewer & Day, 2005, for a study on the confidence–accuracy relationship in a target-present lineup), we strongly encourage conducting more experiments to further investigate the relationship before firm recommendations for practice can be drawn.

The current study echoes previous calls to include adolescents in developmental studies on eyewitness memory (Fitzgerald & Price, 2015; Jack et al., 2013). We embraced this call by studying adolescent's vulnerability to misidentify a bystander from a lineup. Adolescents showed a unique result pattern such that they were more likely than younger, older children, and adults to erroneously identify a bystander as the culprit. The selection of an innocent suspected of a crime may lead to miscarriages of justice. Selecting a bystander is a particularly harmful misidentification error because the bystander was present at the crime scene and will therefore be unable to disprove his or her involvement in the crime with an alibi that links him or her to a different location. An innocent bystander identification is therefore a misidentification error that is difficult to uncover. Our results, that adolescents are most likely to make these innocent bystander identifications, highlight the critical importance of studying adolescents in the eyewitness identification context. More broadly, continued efforts are needed to disentangle how adolescents, in comparison to other age groups, perform in forensically relevant situations.

References

Blank, H., & Launay, C. (2014). How to protect eyewitness memory against the misinformation effect: A meta-analysis of post-warning studies. *Journal of Applied Research in Memory and Cognition*, *3*, 77–88. https://doi.org/10.1016/j.jarmac.2014.03.005

Brainerd, C. J. (2013). Developmental reversals in false memory: A new look at the reliability of children's evidence. *Current Directions in Psychological Science*, *22*, 335–341. https://doi.org/10.1177/0963721413484468

Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin*, *134*, 343–382. https://doi.org/10.1037/0033-2909.134.3.343

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

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil

similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied*, *12*, 11–30. https://doi.org/10.1037/1076-898X.12.1.11

Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology*, *50*, 419–439. https://doi.org/10.1146/annurev.psych.50.1.419

Deffenbacher, K. A., Bornstein, B. H., & Penrod, S. D. (2006). Mugshot exposure effects: Retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law and Human Behavior*, *30*, 287–307. https://doi.org/10.1007/s10979-006-9008-1

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior*, *28*, 687–706. https://doi.org/10.1007/s10979-004-0565-x

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition*, *9*, 288–307. https://doi.org/10.1006/ccog.2000.0447

Fitzgerald, R. J., & Price, H. L. (2015). Eyewitness identification across the life span: A meta-analysis of age differences. *Psychological Bulletin, 141*, 1228–1265. https://doi.org/10.1037/bul0000013

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

Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental study. *Developmental Psychology*, *41*, 625–635. https://doi.org/10.1037/0012-1649.41.4.625

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

Howe, M. L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009). An associative-activation theory of children's and adults' memory illusions. *Journal of Memory and Language*, *60*, 229–251. https://doi.org/10.1016/j.jml.2008.10.002

Jack, F., Leov, J., & Zajac, R. (2013). Age-related differences in the free-recall accounts of child, adolescent, and adult witnesses. *Applied Cognitive Psychology, 28*, 30–38. https://doi.org/10.1002/acp.2951

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

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

Köhnken, G., Milne, R., Memon, A., & Bull, R. (1999). The cognitive interview: A meta-analysis. *Psychology, Crime & Law*, *5*, 3–27. https://doi.org/10.1080/10683169908414991

Lindsay, R. C., 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/Revue Canadienne Des Sciences Du Comportement*, *19*, 463–478. https://doi.org/10.1037/h0079998

Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development*, *76*, 697–712. https://doi.org/10.1111/j.1467-8624.2005.00872.x

McGuire, K., London, K., & Wright, D. B. (2015). Developmental trends in false memory across adolescence and young adulthood: A comparison of DRM and memory conformity paradigms. *Applied Cognitive Psychology*, *29*, 334–344. https://doi.org/10.1002/acp.3114

Memon, A., Meissner, C. A., & Fraser, J. (2010). The Cognitive Interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law*, *16*, 340–372. https://doi.org/10.1037/a0020518

Otgaar, H., Howe, M. L., Brackmann, N., & Smeets, T. (2016). The malleability of developmental trends in neutral and negative memory illusions. *Journal of Experimental Psychology: General*, *145*, 31–55. https://doi.org/10.1037/xge0000127

Otgaar, H., Howe, M. L., Peters, M., Smeets, T., & Moritz, S. (2014). The production of spontaneous false memories across childhood. *Journal of Experimental Child Psychology, 121*, 28–41. https://doi.org/10.1016/j.jecp.2013.11.019

Pozzulo, J. D., & Lindsay, R. C. L. (1997). Increasing correct identifications by children. *Expert Evidence, 5*, 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*, 549–570. https://doi.org/10.1023/A:1025739514042

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

Quas, J. A., Rush, E. B., Yim, I. S., Edelstein, R. S., Otgaar, H., & Smeets, T. (2016). Stress and emotional valence effects on children's versus adolescents' true and false memory. *Memory, 24*, 696–707. https://doi.org/10.1080/09658211.2015.1045909

Read, J. D., Tollestrup, P., Hammersley, R., McFadzen, E., & Christensen, A. (1990). The unconscious transference effect: Are innocent bystanders ever misidentified? *Applied Cognitive Psychology, 4*, 3–31. https://doi.org/10.1002/acp.2350040103

Ross, D. F., Ceci, S. J., Dunning, D., & Toglia, M. P. (1994). Unconscious transference and mistaken identity: When a witness misidentifies a familiar with innocent person. *Journal of Applied Psychology, 79*, 918–930. https://doi.org/10.1037/0021-9010.79.6.918

Ross, D. F., Marsil, D. F., Benton, T. R., Hoffman, R., Warren, A. R., Lindsay, R. C. L., & Metzger, R. (2006). Children's susceptibility to misidentifying a familiar bystander from a lineup: When younger is better. *Law and Human Behavior, 30*, 249–257. https://doi.org/10.1007/s10979-006-9034-z

Sauerland, M., Brackmann, N., & Otgaar, H. (in press). Rapport: little effect on children's, adolescents', and adults' statement quantity, accuracy, and suggestibility. *Journal of Child Custody*. https://doi.org/10.1080/15379418.2018.1509759

Sauerland, M., Raymaekers, L. H. C., Otgaar, H., Memon, A., Waltjen, T. T., Nivo, M., . . . Smeets, T. (2016). Stress, stress-induced cortisol responses, and eyewitness identification performance. *Behavioral Sciences & the Law, 34*, 580–594. https://doi.org/10.1002/bsl.2249

Sauerland, M., & Sporer, S. L. (2009). Fast and confident: Postdicting eyewitness identification accuracy in a field study. *Journal of Experimental Psychology: Applied*, *15*, 46–62. https://doi.org/10.1037/a0014560

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, *100*, 139–156. https://doi.org/10.1037/0033-2909.100.2.139

Shulman, E. P., Smith, A. R., Silva, K., Icenogle, G., Duell, N., Chein, J., & Steinberg, L. (2016). The dual systems model: Review, reappraisal, and reaffirmation. *Developmental Cognitive Neuroscience*, *17*, 103–117. https://doi.org/10.1016/j.dcn.2015.12.010

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence–accuracy relation in eyewitness identification studies. *Psychological Bulletin*, *118*, 315–327. https://doi.org/10.1037/0033-2909.118.3.315

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, *25*, 459–473. https://doi.org/10.1023/A:1012888715007

Stevens, M. C., Kiehl, K. A., Pearlson, G. D., & Calhoun, V. D. (2007). Functional neural networks underlying response inhibition in adolescents and adults. *Behavioural Brain Research*, *181*, 12–22. https://doi.org/10.1016/j.bbr.2007.03.023

Tredoux, C. G. (1998). Statistical inference on measures of lineup fairness. *Law and Human Behavior*, *22*, 217–237. https://doi.org/10.1023/A:1025746220886

Tredoux, C. (1999). Statistical considerations when determining measures of lineup size and lineup bias. *Applied Cognitive Psychology*, *13*, 9–26. https://doi.org/10.1002/(SICI)1099-0720

van Duijvenvoorde, A. C., Peters, S., Braams, B. R., & Crone, E. A. (2016). What motivates adolescents? Neural responses to rewards and their influence on adolescents' risk taking, learning, and cognitive control. *Neuroscience & Biobehavioral Reviews*, *70*, 135–147. https://doi.org/10.1016/j.neubiorev.2016.06.037

van den Bos, W., & Hertwig, R. (2017). Adolescents display distinctive tolerance to ambiguity and to uncertainty during risky decision making. *Scientific Reports, 7*, 1-11. https://doi.org/10.1038/srep40962

Weber, N., & Brewer, N. (2004). Confidence–accuracy calibration in absolute and relative face recognition judgments. *Journal of Experimental Psychology: Applied*, *10*, 156–172. https://doi.org/10.1037/1076-898X.10.3.156

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, *22*, 603–647. https://doi.org/10.1023/A:1025750605807

Wixted, J. T., Read, J. D., & Lindsay, D. S. (2016). The effect of retention interval on the eyewitness identification confidence–accuracy relationship. *Journal of Applied Research in Memory and Cognition*, *5*, 192–203. https://doi.org/10.1016/j.jarmac.2016.04.006