Eyewitness Identification across the Life Span: A Meta-Analysis of Age Differences

Ryan J. Fitzgerald* and Heather L. Price

Department of Psychology

University of Regina

This article is published in *Psychological Bulletin*.

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

Abstract

Lineup identifications are often a critical component of criminal investigations. Over the past 35 years, researchers have been conducting empirical studies to assess the impact of witness age on identification accuracy. A previous meta-analysis indicated that children are less likely than adults to correctly reject a lineup that does not contain the culprit, but children 5 years and older are as likely as adults to make a correct identification if the culprit is in the lineup (Pozzulo & Lindsay, 1998). We report an updated meta-analysis of age differences in eyewitness identification, summarizing data from 20,244 participants across 91 studies. Contrary to extant reviews, we adopt a life span approach and examine witnesses from early childhood to late adulthood. Children's increased tendency to erroneously select a culprit-absent lineup member was replicated. Children were also less likely than young adults to correctly identify the culprit. Group data from culprit-absent and culprit-present lineups were used to produce signal detection measures, which indicated young adults were better able than children to discriminate between guilty and innocent suspects. A strikingly similar pattern emerged for older adults, who had even stronger deficits in discriminability than children, relative to adults. Although identifications by young adults were the most reliable, identifications by all witnesses had probative value.

Eyewitness Identification across the Life Span: A Meta-Analysis of Age Differences

The ability to recognize faces is crucial to social interaction. From infancy, we are continuously exposed to faces that comprise our most critical environmental stimuli, and facial recognition improves drastically during the early stages of development (Nelson, 2001). Deficits in developing these skills can contribute to substantial social impairment (e.g., autism spectrum disorder, prosopagnosia), as social norms dictate adequate recognition of previously encountered faces. The extant research has made clear that face recognition is a specialized skill involving unique brain regions that is distinct from the ability to recognize other objects (Richler & Gauthier, 2014; Schwartz, 2014).

The ubiquity of face recognition in daily interactions makes understanding this ability of widespread interest. However, remembering a face is particularly crucial when recognition of a stranger is required to identify the perpetrator of a crime. Crimes are often committed in the presence of others, but if the perpetrator is unknown to the witness, difficulty in later recognizing the perpetrator can thwart criminal investigations. There is an intuitive sense that if a person was observed committing a crime, identifying the criminal should not be difficult. However, decades of research on eyewitness identification abilities make clear that this intuition is flawed.

To date, DNA evidence has revealed more than 300 cases of wrongful conviction in the United States alone. In approximately 70% of those cases, eyewitness identification errors were a contributing factor (Innocence Project, 2014). At around the same time that legal inquiries into wrongful convictions were revealing an alarming number of confirmed false identifications (Brooks, 1983; Devlin, 1976), psychologist Gary Wells (1978) published a landmark article that partitioned the influences on eyewitness identification accuracy into variables that can be controlled through investigative policies (system variables) and variables that are beyond the control of the justice system (estimator variables). The experimental work that flowed from this framework has revealed weaknesses not only in the memories of eyewitnesses, but also in the methods through which legal investigators extract those memories. For instance, research on system variables has demonstrated an increased risk of false identification when witnesses are not warned of the culprit's potential absence from the lineup (Malpass & Devine, 1981), when lineup members are presented simultaneously (Lindsay & Wells, 1985), and when lineup members do not all match a description of the culprit (Wells, Rydell, & Seelau, 1993).

Guidelines for assessing the reliability of eyewitness identification have been set out in influential legal decisions (*Manson v. Braithwaite*, 1977; *Neil v. Biggers*, 1972; *R v. Turnbull and others*, 1976). The guidelines advise consideration primarily of information processing factors. Thus, in spite of the empirically demonstrated associations between system variables and false identification, estimator variables appear to carry the most weight in courtrooms. For instance, in the United States, jurors are instructed to consider the opportunity to view the culprit, the attention paid to the culprit, the quality of the prelineup description of the culprit, the retention interval between the event and the identification, and, finally, the degree of confidence expressed by the witness (*Neil v. Biggers*, 1972).

The present meta-analytic review focuses on another key estimator variable, the age of the witness. Early quantitative reviews of age effects on eyewitness identification indicated that both children and older adults are as likely as young adults to identify the culprit when the person they are trying to identify is in the lineup, but young adults are more likely than their younger and older counterparts to correctly reject the lineup when the culprit is absent (Bartlett & Memon, 2007; Pozzulo & Lindsay, 1998). However, practical and methodological limitations that were characteristic of the data summarized in these early reviews give reason to question the reliability of these findings. Although many of these problems were addressed in a recent meta-analytic comparison specifically assessing the performance of older witnesses (Sporer & Martschuk, 2014), this review was limited by its exclusion of witnesses under 16 years of age.

Our meta-analytic review is the first to examine age differences in eyewitness identification across the life span. Though behavioral observations suggest that many, perhaps most, broad cognitive processes developing throughout infancy and childhood are similar to those that decline with age (e.g., McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002; Salthouse, 2004), relatively little attention has been paid to marrying the developing cognitive processes of childhood with the declining cognitive processes of aging. Indeed, despite the very similar patterns of identification data observed between young children and older adult witnesses, in our search for studies directly comparing these age groups we found only one. This article represents our attempt to bridge this gap. We have extracted data from over 20,000 participants, yielding a rich dataset that allows for a comprehensive understanding of age differences in eyewitness identification outcomes. Although some of our conclusions represent a departure from existing views in the literature, we argue that this new perspective provides a more parsimonious and intuitive understanding of age effects on eyewitness identification.

Children as Witnesses

More than 15 years ago, Pozzulo and Lindsay (1998) meta-analyzed age differences in identification accuracy. As the only review and synthesis of developmental differences in eyewitness identification, the field has relied heavily upon the reported findings when developing new theory and methods. However, since that 1998 publication, a large volume of studies have been conducted and whether or not the previous meta-analytic findings will stand the test of time is unclear. In the literature review that follows, we focus on two key findings from the Pozzulo and Lindsay meta-analysis: (a) children are less likely than young adults to reject culprit-absent lineups, and (b) children "can" identify the culprit as effectively as young adults.

Children Are Less Likely Than Young Adults to Reject Culprit-Absent Lineups

Without question, the most influential finding in the child eyewitness identification literature is that children have an increased propensity to choose innocent lineup members from culprit-absent lineups. The earliest studies on children's eyewitness identification did not include culprit-absent lineups (Dent & Stephenson, 1979; Marin, Holmes, Guth, & Kovac, 1979), and the earliest child witness studies that included culprit-absent lineups did not include adult comparisons groups (Peters, 1987; Yarmey, 1988), so awareness of children's heightened tendency to choose did not emerge in the literature until 10 years after the first explorations of children's eyewitness identification (Parker & Carranza, 1989). Culprit-absent lineups were routinely administered in subsequent investigations. In 1998, Pozzulo and Lindsay meta-analyzed nine studies that compared children's and adults' culprit-absent lineup responses. Correct rejection rates were significantly higher for adults than for three age groups of children (4, 7–8, and 12–13 years) and numerically higher than for the only other child comparison group (5– 6 years), leading Pozzulo and Lindsay to conclude that children were less proficient than adults on culprit-absent tasks. In the 15+ years since their meta-analysis, this effect has been replicated on numerous occasions.

Several theories have been advanced to explain children's difficulty with culprit-absent lineups. One explanation focuses on the implicit social demands of a lineup task. Children are sensitive to suggestive interviewing techniques and tend to give the answer they think the interviewer wants (Bruck & Ceci, 1999; Ceci & Bruck, 1993). A lineup task has been considered analogous to a leading question (Davies, 1996). The mere presentation of photographs by an authority figure may lead children to presume that the target person is present and that their objective is to pick the “right” photograph (Lowenstein, Blank, & Sauer, 2010; Parker & Carranza, 1989). Providing children with a pre-lineup admonition about the possibility of the culprit’s absence has been shown to reduce children’s false identification rate, but even when such instructions are given children remain less likely than adults to correctly reject culprit-absent lineups (Keast, Brewer, & Wells, 2007; Pozzulo & Dempsey, 2006).

The cognitive skills required to recognize when the appropriate response is to reject a lineup may simply be underdeveloped in children. Culprit-absent lineups can present a rather difficult task. If the culprit is present, a correct decision can typically be made via familiarity processes. That is, the witness can correctly identify the person without recalling any contextual details of where and when the person was encountered. By contrast, a correct decision on a culprit-absent lineup requires a recall-then-reject strategy (Gross & Hayne, 1996). First, the culprit must be recalled and this representation must be compared with each of the lineup members in search of a match. Given that the processes required to recall information are believed to develop later than the processes required to recognize a previously encountered stimulus (Kail, 1990), the difficulty of a culprit-absent lineup may pose too great of a task for the developing mind of some children.

Developmental differences in response inhibition— deliberately withholding an inappropriate response— could also be involved. Sinopoli, Schachar, and Dennis (2011) recently observed age-related increases in both cancellation and restraint inhibition among adolescents relative to children, and others have suggested response inhibition continues to develop until adulthood (e.g., Nigg, 2000). Thus, children's inclination to erroneously choose innocent lineup members could be related to an inability to stop themselves from positively responding to the lineup task. Consistent with this notion, children typically reject more lineups when they can do so by choosing a visual representation of the culprit's absence (e.g., Davies, Tarrant, & Flin, 1989; Zajac & Karageorge, 2009). This effect has been explained as a consequence of making the process of rejecting a lineup more similar to the process of selecting a lineup member.

Some researchers have proposed that age differences in responding to culprit-absent lineups reflect children's use of less effective face processing strategies (Davies et al., 1989). Early face recognition research led to the suggestion that young children rely primarily on an ineffective feature-based encoding strategy, which involves encoding specific facial features in a piecemeal fashion (Diamond & Carey, 1977). Using such a strategy, children could be expected to select any lineup member possessing the features that were encoded from the culprit's face. For instance, a child who noticed that the culprit had a large nose might select any lineup member with a large nose. Such an error would seem less likely for adults, who are known to process faces holistically. Although there is some debate about the precise nature of holistic processing, the strategy is generally considered to involve viewing the face as a whole and taking the spacing of features into account (Tanaka & Farah, 1993). Although holistic processing deficits have been observed in children (Carey, Diamond, & Woods, 1980; Mondloch, Le Grand, & Maurer, 2002), a recent review of the face processing literature led the authors to conclude that holistic processing is fully mature by about age 5 and that age differences in face recognition performance reflect differences in general cognitive factors, such as memory or attention (Crookes & McKone, 2009).

Our literature review revealed several explanations for children's increased proclivity to choose on culprit-absent lineup tasks. Children appear less aware than young adults of the option to choose none of the lineup members. Even when children understand that they can reject the lineup, they still tend to make more errors than young adults. An underdeveloped ability to recall the previously viewed face and an inability to refrain from making inappropriate responses have also been implicated as contributing factors to children's high rate of false positive responding. Given that a child's decision to identify a lineup member almost certainly involves a combination of factors, it is difficult to tease apart the relative contributions of the various potential mechanisms involved.

Children "Can" Identify the Culprit as Effectively as Young Adults

It has become standard for researchers to inform their readers early in child eyewitness identification articles that although children are less likely than young adults to reject culprit-absent lineups, if the culprit is in the lineup children aged 5 years and older make correct identifications at rates similar to those for young adults. To someone unfamiliar with this research, we imagine the latter claim would be surprising. Nevertheless, the consensus among child witness researchers is remarkable. In the past few years alone, strong claims have been made about children's correct identifications from culprit-present lineups: Havard and Memon (2013) state that "[i]t is well established that children (as young as 5 years) can correctly identify a culprit from a target present (TP) line-up as accurately as adults" (p. 50); Humphries, Holliday, and Flowe (2012) write that "children (5 years and over) are as likely as adults to correctly identify the culprit when shown a target-present lineup" (p. 149); and Dunlevy and Cherryman (2013) comment that "children aged 5–14 years show performance comparable with that of adults for correct identifications in target-present line-ups" (p. 285). We found more than a dozen articles containing similar statements. More often than not, the statement is supported with a citation of Pozzulo and Lindsay's (1998) meta-analysis.

Pozzulo and Lindsay (1998) found that all children were less likely than adults to correctly reject a culprit-absent lineup, but only the very young children (4 years old) were less likely than adults to correctly identify a culprit. In fact, Pozzulo and Lindsay reported a significantly higher correct identification rate for the 5–6 years group than for the adult group. We previously noted that face perception may be fully mature by about age 5 (Crookes & McKone, 2009). This could be considered to correspond well with the finding that children aged 5 and older are as likely as adults to make a correct identification. Although numerous studies have revealed age-related increases in correct responses to laboratory-style face recognition tasks (Blaney & Winograd, 1978; Chance, Turner, & Goldstein, 1982; Chung & Thomson, 1995; Ellis & Flin, 1990; Shapiro & Penrod, 1986), Crooks and McKone argue that such differences reflect developmental advances in general cognition (e.g., memory, attention) rather than advances in face perception per se.

Although research has demonstrated that children can achieve similar correct identification rates to those of adults, these reports of equivalent *performance* may not be indicative of equivalent *ability*. Although child eyewitness identification researchers have generally steered clear of this distinction, their language suggests skepticism about children's correct identification abilities. Specifically, rather than stating that children *are* as good as adults on culprit-present lineup tasks, researchers commonly note that children *can* identify the culprit as effectively as adults.

We suspect that reports of comparable correct identification rates between children and adults are merely indicative of children's increased willingness to guess in the face of uncertainty. There is a consensus in the literature that children are more likely than adults to select a lineup member when the culprit is absent, but it seems unlikely that this increased tendency to choose would be limited to culprit-absent lineups. Rather, children's weakness on culprit-absent lineups and their relative strength on culprit-present lineups is likely indicative of a more general proclivity to choose or, as it is known in the basic cognitive literature, a liberal response bias. Response bias is often calculated using measures derived from signal detection theory, which distinguishes between response bias and sensitivity (Green & Swets, 1966; Macmillan & Creelman, 1991). Our position is that children are not as sensitive as adults to the presence of the culprit, but age-related differences in the ability to correctly identify the culprit tend to be obscured by children's more liberal response bias.

This response bias account of children's correct identification rates is not new. In early studies, researchers attributed children's strong culprit-present performance to their adoption of a less-stringent threshold for making an identification (Dekle, Beal, Elliott, & Huneycutt, 1996; Parker & Carranza, 1989; Parker & Ryan, 1993). However, this possibility has not received much attention in recent years, which is perhaps related to what has become a common practice of not reporting rejection rates for culprit-present lineups (the prevalence of which can be seen in the Appendix, Table A1).

In addition to having doubts about children's *ability* to make correct identifications from culprit-present lineups, we also have reservations about the evidence that children *perform* as well as adults on culprit-present lineups. A careful review of the only meta-analytic comparison between children's and adults' eyewitness identification revealed several issues that may affect the reliability of the conclusions drawn (Pozzulo & Lindsay, 1998). One problem with the meta-analysis was that a notable proportion of the summarized studies lacked methodological rigor. For instance, in one study a large advantage in correct identifications was found for children relative to adults (Dekle et al., 1996); however, the culprit-present lineup data comprised responses from only 18 children and 67 adults. Small samples such as these were common during that era.

Another common methodological artifact of pre-1998 studies was the use of forced-choice paradigms (Marin, Holmes, Guth, & Kovac, 1979; Parker, Haverfield, & Baker-Thomas, 1986). Forcing participants to pick a lineup member contrasts starkly with eyewitness identification procedures in the field and almost all contemporary research, in which witnesses have the option to reject all of the lineup members. Even when an option to reject was provided, children did not always make use of it. In one study (Dekle et al., 1996), 94% of children presented with a culprit-present lineup picked one of the lineup members (relative to only 43% of adults). Dekle and colleagues explicitly reported instructing the adults not to guess (i.e., adopt a conservative response bias), but did not report providing the same instruction to the children.

Another major concern with the 1998 meta-analysis is that very few studies examining age differences in eyewitness identification were available at that time. A few problematic primary studies might not have much influence on a meta-analysis that summarizes a well-populated literature. However, Pozzulo and Lindsay's (1998) meta-analysis summarized data from only 15 primary studies and the review required multiple meta-analyses to compare adults with four age groups of children. As a consequence, most of the meta-analyses only summarized three to five primary studies. In addition, two studies that showed an adult advantage in correct identifications were not included in the 1998 meta-analysis (Mertin, 1989; Yarmey, 1988). Given the small number of studies that were summarized, including these studies might have had a substantial impact on the difference between children and adults.

Older Adults as Witnesses

Though there has been a recent modest increase in research interest in older adult eyewitnesses, this population has generally received substantially less attention than have child witnesses. The lack of empirical investigation of older adult witnesses is characteristic of the forensic psychology field more generally and has recently been the subject of calls for additional research (Brank, 2007). In the research that has been conducted with older adults, identification patterns have been similar to those of children. Specifically, relative to young adults, older adults have evinced higher false identification rates in culprit-absent lineups (e.g., Memon, Hope, Bartlett, & Bull, 2002), with similar levels of correct identifications in culprit-present lineups (e.g., Memon & Bartlett, 2002).

The existence of these parallel literatures should not be surprising given the mirrored cognitive abilities (e.g., memory capacity, response inhibition, attention) and susceptibility to social demands that have been noted across the life span (e.g., Salthouse, 2004). However, despite the similarities observed in response patterns, both the eyewitness identification literature and basic theoretical memory literature have focused on different mechanistic explanations. In general, whereas the child eyewitness literature has focused on both social and cognitive factors, the older adult eyewitiness literature has primarily focused on cognitive factors, with more limited attention to social factors.

The findings in the older adult eyewitness literature are more easily compiled and understood than for child eyewitnesses due to two recent reviews and a meta-analysis. In 2007, Bartlett and Memon examined 19 datasets from 12 lineup studies and established that older adults showed a key similarity to child witnesses: low correct rejection rates on culprit-absent lineup tasks. Overall, the average correct identification rate for older adults was only 4% lower than for young adults. However, the average correct rejection rate was 28% lower for older adults relative to young adults. Thus, as with child witnesses, Bartlett and Memon's findings clearly indicated that although older adults were not always less likely to correctly identify the target, they were consistently less likely to reject culprit-absent lineups. In a subsequent review, Bartlett (2014) confirmed the initial report of an age-related decline in correct rejections and also made note of recent studies demonstrating lower correct identification rates in older relative to younger adults. Of particular relevance for the present review, a recently conducted meta-analysis of adult age differences in eyewitness identification revealed better performance for young adults relative to older adults regardless of whether the lineup contained the culprit or not (Sporer & Martschuk, 2014). The meta-analysis also indicated that, relative to young adults, the odds that a lineup member would be chosen were 2.3 times greater for older adults.

Consistent with the child witness literature, social and cognitive mechanisms have been suggested as factors contributing to older adults' increased propensity to choose. However, although some attention has been paid to consideration of social factors for older adult witnesses, it has certainly been less so than for children and the specific nature of the social influence proposed has differed across age groups. While a substantive focus in research with children has been on compliance with adult authority figures and their responsiveness to situational demands as novice learners (e.g., Ceci & Bruck, 1993), it has been suggested that older adults may be especially motivated to "help" with police investigations by choosing from a lineup

(Gallagher, Maguire, Mastrofski, & Reisig, 2001; Sporer & Martschuk, 2014). There has been little other exploration of social factors that may contribute to older adults' performance.

Because children are inherently attuned to adults' social cues, controlling the perception of implicit situational demands has been a substantive focus for child witness researchers. Many researchers have worked to reduce children's apparent belief that they "should" make a choice via increasing the salience of rejection options. The inclusion of a salient rejection option within a lineup aims to increase attention to the possibility of "choosing not to choose" (Zajac & Karageorge, 2009). Although these efforts have shown some success with children, salient rejection options have only recently been examined with older adults and results have been mixed (Gentle, 2012; Havard, n.d.).

There is also concern about older adults' memory for the target. Although age-related deficits in memory tasks are quite consistently observed in older adults, these effects can be magnified or minimized under certain conditions. For instance, age-related deficits tend to increase with controlled or conscious versus automatic processing (Balota, Dolan, & Duchek, 2000) and decrease with environmental or contextual support (Anderson & Craik, 2000). Several researchers have explored the utility of contextual support, given older adults' documented difficulty with encoding context (Spencer & Raz, 1995) and source memory (e.g., Aizpurua, Garcia-Bajos, & Migueles, 2011; Bornstein, 1995; Johnson, Hashtroudi, & Lindsay, 1993). Though the benefits of contextual support have been inconsistent, some types of context reinstatement or pre-identification practice trials have reduced older adults' misidentification rates (Wilcock & Bull, 2010; Wilcock, Bull, & Vrij, 2007; but see Memon, Hope, Bartlett, & Bull, 2002; Searcy, Bartlett, Memon, & Swanson, 2001).

In general, most interventions for older witnesses have been focused on increasing the likelihood that older adults will adopt a stricter decision criterion. Sequential presentation of lineup members is perhaps the most well-known innovation in lineup identification research that has promoted the use of a strict decision criterion. Although the sequential lineup has been found to reduce older adults' misidentification rates, it has also been found to reduce correct identification rates (e.g., Memon & Gabbert, 2003a, 2003b). Sporer and Martschuk's (2014) meta-analysis provided further evidence for use of a stricter decision criterion in sequential lineups. However, the authors were cautious in their conclusions about the use of sequential lineups with older adults, calling for a critical evaluation of the procedure. In sum, older adults perform very similarly to children on lineup identification tasks and the explanations for these patterns have focused on similar cognitive processes and somewhat different social processes.

Memory Processes across the Life Span

Although life span memory theory has not been directly applied to eyewitness identifications, there are emerging theories of memory that predict dissociative processes during cognitive development and cognitive aging. The evidence supporting distinct memory processes in children and older adults led Craik and Bialystok (2006) to argue that aging is not simply "development in reverse." There is a growing body of work supporting the idea that cognitive processes that share surface characteristics are driven by basic differences in children's and older adults' memory processes. For example, differences between older adults and children have been observed in working memory (Cowan, Naveh-Benjamin, Kilb, & Saults, 2006; Sander, Werkle-Bergner, & Lindenberger, 2011), top-down control and binding to working memory (Sander, Lindenberger, & Werkle-Bergner, 2012), and the ability to hold and retrieve information over the short-term (Fandakova, Sander, Werkle-Bergner, & Shing, 2014).

In the past 10 years, there has been a steady increase in calls for theoretical models that describe memory mechanisms over the life span. Relying on a dual-process account of memory that distinguishes between associative (i.e., automatic, binding processes of memory) and strategic (i.e., effortful control processes that assist with encoding and retrieval) components of episodic memory, Shing et al. (2010, 2008) dissociated children's and older adults' memory processes. Developmental changes in children's memory have been traced to the associative components and the development of the prefrontal cortex. In older adults, a decline in both the associative and strategic components of memory has been linked to changes in the prefrontal cortex and the medial temporal lobe (Shing et al., 2010; Shing, Werkle-Bergner, Li, & Lindenberger, 2008). Shing and colleagues proposed that though the strategic component is lower in both children and older adults relative to younger adults, the associative component is impoverished in older adults relative to both children and young adults, who do not differ so dramatically (Cowan et al., 2006). Thus, associative memory is a potential source of differences between children's and older adults' episodic memory performance.

Shing et al. (2010, 2008) noted the similarities between their dual-process account and Jacoby's (1991) process dissociation model. Jacoby distinguished between recollection, which is a controlled retrieval of detailed contextual information, and familiarity, which is an automatic feeling of knowing without specific remembrance. Relative to young adults, older adults have been found to rely relatively more on familiarity processes during recognition tasks due to impairment in recollection (Healy, Light, & Chung, 2005). Reliance on familiarity is likely to result in a lower criterion threshold and higher rates of choosing (Yonelinas, 2002). Familiarity (and associative) processes mature early in childhood, whereas recollection (and strategic) processes develop throughout middle childhood (Brainerd & Reyna, 2004; Shing et al., 2010, 2008). Thus, both older adults and children may be more likely than young adults to rely on familiarity processes, with older adults also more likely to be disadvantaged due to weaker associative processes.

The Present Meta-Analysis

Given the above evidence that older adults and children seem to perform less accurately on lineup tasks than do young adults, we were interested in better understanding the differential response patterns across the life span. Basic cognitive research suggests that behavioral similarities in memory tasks between children and older adults may be driven by different mechanisms (Craik & Bialystok, 2006). Such conclusions raise questions about the relative contribution of memory on age differences in eyewitness identification decisions. Although a combination of social and cognitive factors likely contribute to any identification decision, the present research aims to bring together basic cognitive and life span developmental research with the applied eyewitness literature to better understand how face recognition decisions differ across the life span. Exploration of how distinct types of memory may vary across the life span will also contribute to much needed integrative life span theories of cognition (Sander et al., 2012).

Thus far, we have discussed age effects on eyewitness identification separately for culprit-present and culprit-absent lineups. However, a composite measure that takes into account responses on both lineup types provides a more comprehensive understanding of age effects on eyewitness identification. The traditional measure for examining performance across culprit-present and culprit-absent lineups is the diagnosticity ratio, which is calculated by dividing the guilty suspect identification rate (correct identification) by the innocent suspect misidentification rate (false identification). However, the diagnosticity ratio has been criticized for its susceptibility to influences on response criterion (Mickes, Moreland, Clark, & Wixted, 2014; Wixted & Mickes, 2012, 2014). Specifically, as a procedure becomes increasingly conservative, the diagnosticity ratio will generally also tend to increase. As an alternative to the diagnosticity ratio, Mickes et al. (2014) recommend a measure of sensitivity (*d'*) derived from signal detection theory. Sensitivity analyses were initially proposed for measuring discrimination between signals and noise (Green & Swets, 1966); however, signal detection theory can be applied to any experiment that tests discrimination between two types of stimuli (Macmillan & Creelman, 1991; Stanislaw & Todorov, 1999). Over the past several decades, signal detection theory has been used widely in recognition memory studies that require participants to learn a list of items at study and discriminate between old and new items at test.

Although *d'* has a lengthy history in recognition memory experiments, its application in eyewitness identification experiments is a recent development. Signal detection theory was designed for analyzing yes/no response data, collected over multiple trials. Applying signal detection theory to list-learning experiments is generally straightforward because a *d'* value can be calculated for each participant and group level variance can be calculated for computation of commonly used inferential statistics. However, signal detection theory was not designed for analyzing data obtained from eyewitness identification experiments, which typically expose participants to a single culprit and then test recognition with a single lineup that may or may not contain the culprit. Eyewitness identification researchers have developed paradigms for exposing participants to numerous targets and administering lineups across multiple trials (e.g., Meissner, Tredoux, Parker, & MacLin, 2005); however, experiments employing such paradigms represent only a small minority of the eyewitness identification literature. For most eyewitness identification experiments, *d'* can only be computed at the group level. This can be achieved by treating the proportion of guilty suspect identifications in the culprit-present condition as the hit rate and treating the proportion of innocent suspect misidentifications in the culprit-absent condition as the false alarm rate (Clark, 2012). Thus, *d'* can be computed to represent the extent to which eyewitnesses can discriminate between guilty and innocent suspects on lineup tasks.

The present meta-analytic review examines age differences in eyewitness identification. In addition to examining age differences in *d'*, we examined two measures of response bias, one that represents the inclination to identify the suspect (c_{suspect}) and one that represents a more general inclination to identify any of the lineup members (choosing). We predicted higher rates of choosing for children and older adults than for young adults, but greater sensitivity for young adults than for children and older adults.

Method

Literature Search

Search procedures. A search was conducted to locate studies comparing two or more age groups on a lineup identification task. Four databases (*PsycINFO*, *Web of Science, Google Scholar*, and *ProQuest Dissertations and Theses*) were searched using various combinations of the following terms: accuracy, adolescent, adult, age, aging, child, develop, eyewitness, face, false, identification, lineup, memory, old, preschool, recognition, testimony, witness, and young. Following a search of the databases, a snowball method was used to examine the reference sections and citation records of relevant articles to locate any additional studies. In addition to searching for unpublished theses, we contacted more than 60 authors who have previously published articles examining age differences in eyewitness identification and requested unpublished work that would meet our inclusion criteria.

Inclusion criteria. For analyses of the main effect of age on eyewitness identification, a study required the following characteristics: (a) two or more age groups were compared on an eyewitness identification task; (b) the event was experienced via a live interaction or a video/slideshow presentation (studies were excluded if they employed laboratory-style face recognition paradigms that consisted of viewing and testing memory for a series of still photographs); (c) memory was tested via an identification test containing multiple lineup members, rather than a single person (single-person memory tests, or showups, were analyzed separately in moderator analyses); (d) researchers tested recognition memory using a lineup containing one previously encountered person (culprit-present lineup) or zero previously encountered persons (culprit-absent lineup); an exception to this rule was for lineups containing an innocent bystander, which were included and treated as culprit-absent lineups; (e) participants made a discrete, categorical lineup decision (i.e., they identified a lineup member or they rejected all lineup members; as opposed to an exclusive confidence rating about the relative likelihood that a lineup member was the culprit); (f) researchers reported culprit-present and culprit-absent performance separately, as opposed to an overall accuracy rate with culprit-present and culprit-absent conditions collapsed; and (g) researchers reported sufficient information to compute an odds ratio.

Characteristics of the final dataset. The search, which concluded in October 2014, produced 85 published journal articles, book chapters, or unpublished manuscripts/theses containing at least one study that met the inclusion criteria. Publication dates ranged between 1979 and 2014. Some articles had multiple studies that met the inclusion criteria. In total, data from 20,280 participants were extracted from 91 studies (69 published, 22 unpublished). A subset of data from two studies was excluded because the age group was notably younger (M < 36 months; Cain, Baker-Ward, & Eaton, 2005) or notably older (range = 75–94 years; Scogin, Calhoon, & D'Errico, 1994) than all other groups in the literature. After these exclusions, data from 20,244 participants remained. The studies were organized into two datasets. The first dataset comprised 60 studies that compared young adults with child witness (29 published, five unpublished), older adult witnesses (21 published, four unpublished), or both (zero published, one unpublished). The second dataset comprised 42 studies that compared different age groupings of child witnesses (28 published, 14 unpublished). Some studies contributed data to both datasets because they included comparisons between young adult and child witnesses as well as comparisons between two or more child groups. The lineup response rates for the primary studies in both datasets are provided in Appendix A.

Meta-Analytic Procedures

Robust variance estimation. Not all effect sizes in the datasets were independent. Hedges, Tipton, and Johnson (2010) describe several types of effect size dependence. One type of dependency involves shared comparison groups. For example, Humphries and colleagues compared 5- to 6-year-olds, 9- to 10- year-olds, and young adults (Humphries et al., 2012). For a meta-analysis comparing young adults and children, two effect sizes could be computed: (a) young adults versus 5- to 6-year-olds, and (b) young adults versus 9- to 10-year-olds. These effect sizes would be dependent because they have a comparison group in common (i.e., young adults). A second type of dependency occurs when multiple effect sizes are nested within a study. For example, Pozzulo and colleagues manipulated lineup procedure (simultaneous vs.

sequential vs. elimination) for adolescent and young adult participants (Pozzulo, Dempsey, & Pettalia, 2013). Given that lineup procedure was manipulated between-subjects, an effect size for age differences in the response rates could be calculated for each of the three lineup procedures. Although these effect sizes would not have any participants in common, dependence would nevertheless be present because of commonalities in the experimental procedure across conditions (e.g., the identity of the lineup members was constant across conditions). A third type of dependence, referred to as hierarchical dependence, occurs when multiple studies are conducted by the same research group. For example, more than 10 studies in the current dataset were conducted in the Pozzulo lab. Again, the effect sizes for these studies might not have any participants in common, but a research group could be expected to use similar procedures across studies (e.g., participants may be sampled from the same pool).

Tanner-Smith and Tipton (2014) note that most meta-analysts have ignored effect size dependencies and treated the effect sizes as though they were independent. If the effect size dependency arises from commonalities in the experimental procedure and the effect sizes are weighted using the fixed-effect model, ignoring the dependency might not be problematic. However, the fixed-effect model is rarely appropriate in the social sciences (Borenstein, Hedges, Higgins, & Rothstein, 2010). If the random-effects model is applied, ignoring the dependencies would compromise the integrity of the meta-analysis by artificially reducing variance estimates (Borenstein et al., 2010) and systematically linking a study's weight to the number of effect sizes it contributes (Scammacca, Roberts, & Stuebing, 2014).

One way to address the dependency is to compute an aggregate effect size for each study and perform the meta-analysis on the average effect sizes for each study, which would then be independent. Although this approach is effective in eliminating the dependency, combining unique effect sizes into an average effect size results in the loss of potentially valuable information (Tipton, 2014). Fortunately, a method known as robust variance estimation (Hedges, Tipton, & Johnson, 2010) was recently developed to address effect size dependence without discarding the unique information provided by multiple effect sizes within a study.

We meta-analyzed age differences using *robumeta*, which is a package for computing robust variance estimation with *R* statistical software (http://www.R-project.org). Robust variance estimation addresses the very types of dependency that are characteristic of the current dataset by making an adjustment to the standard error of each effect size (Hedges et al., 2010; Tanner-Smith & Tipton, 2014). Robust variance estimation is a particularly desirable approach for the present analyses because it can simultaneously address multiple types of dependency (Tanner-Smith & Tipton, 2014) and can be applied for any type of effect size. Until recently, robust variance estimation could only be applied confidently if a meta-analysis summarized at least 10 studies for main effects and at least 40 studies for meta-regression coefficients; however, a small-sample correction is now available (Tipton, 2014). Although the datasets for many of our analyses met the minimum requirements for performing robust variance estimation without the small-sample correction, we followed Tipton's recommendation to apply the small-sample correction for all robust variance estimation analyses.

Weighting method. Hedges et al. (2010) proposed two weighting methods for robust variance estimation, one designed for correlated effects (for effect sizes nested within the same study) and one designed for hierarchical effects (for effect sizes nested within some type of cluster, such as a research laboratory). Both methods are available in the *robumeta* package. We used the correlated effects weighting method, which Tanner-Smith and Tipton (2014) suggest is the most appropriate choice for meta-analyses in which the same group is compared with multiple comparison groups (e.g., young adults vs. 6-year-olds and young adults vs. 10-year-olds). For the correlated effects weighting method, an interclass correlation must be specified. Although the correlation between effects is rarely reported in primary studies, robust variance estimation is generally unaffected by changes in rho and sensitivity analyses can be used to confirm this (Hedges et al., 2010). Accordingly, we set rho to the default setting (.80) and performed sensitivity analyses to ensure changes in rho would not substantially impact our interpretation of the results. In all cases, the sensitivity analyses indicated the specification of rho had negligible effects on the results and sensitivity analyses are not discussed further.

Outliers. The treatment of outliers in meta-analysis has been the subject of considerable debate. If effects that are substantially larger than those typical in the literature are included, the results of a meta-analysis could be distorted (Lipsey & Wilson, 2001). However, some heterogeneity among effect sizes should be expected in a meta-analysis (Higgins, 2008) and removing all outliers may not be desirable (Hedges & Olkin, 1985). Higgins (2008) suggests running the analyses twice, once with outliers excluded and once with outliers included, to determine if the results are robust to the inclusion of outliers. Accordingly, we report all analyses with outliers excluded and make note of the small number of cases in which the significance of the difference was affected by the inclusion of outliers. Outliers were identified though calculation of standardized residuals with the random-effects model using *Comprehensive Meta-Analysis* software (Version 2.0; Borenstein et al., 2005). Outliers were defined as effect sizes with standardized residuals greater than 1.96 (Hedges & Olkin, 1985). An iterative approach was applied in which one outlying effect was removed at a time, until all standardized residuals were below 1.96.

Outcome Measures

We compared age groups on eight outcome measures. Three of the outcomes were responses to culprit-present lineups (hits, filler identifications, and incorrect rejections). One outcome was responses to culprit-absent lineups (correct rejections). The final four outcomes were calculated using data from culprit-present and culprit-absent lineups (choosing, diagnosticity, sensitivity, and response bias).

Hits. The common practice in eyewitness identification research is to expose the witness to a target person, often referred to as the culprit or the perpetrator, and ask the witness if that person is in the lineup. A hit (also known as a culprit/correct identification) occurs if the witness identifies the target person from a lineup. The hit rate was calculated by dividing the number of culprit identifications by the total number of culprit-present lineup responses (culprit identifications + filler identifications + lineups rejections). For effect size calculations, hits were treated as a binary outcome (e.g., the culprit was identified or the culprit was not identified) and odds ratios (*OR*) were calculated for comparing hit rates between two age groups.

Odds are calculated by dividing the number of event occurrences by the number of event non-occurrences. For example, if 20 participants made a hit and 10 did not, the odds of a hit would be 2.00. *OR*s are calculated by dividing the odds of an event occurrence in one group by the odds of an event occurrence in another group. An *OR* of 1.00 indicates perfect unity between two groups in the odds of an outcome. Interpretation of *OR*s above or below unity depends on how the *OR*s are calculated. Specifically, it depends on which odds serve as the numerator and which odds serve as the denominator. Imagine that the odds of a hit are 2.00 for young adults and 0.50 for children. If the odds for young adults were used as the numerator, the *OR* would be 4.00 (2.00/0.50). This would indicate the odds of a hit for young adults are four times greater than the odds for older adults. Conversely, if the odds for children were used as the numerator, the OR would be 0.25 (0.50/2.00). If calculated this way, the *OR* would be interpreted as an indication that the odds of a hit for children are 0.25 times the odds of a hit for young adults. *OR*s below unity are considered to be less intuitive than *OR*s above unity (Osborne, 2006). Therefore, we always used the larger odds as the numerator when calculating *OR*s.

One negative aspect of *OR*s is that the lower end (< 1.00) has a restricted range, whereas the upper end (> 1.00) has no bounds, which produces a skewed distribution (Bland & Altman, 2000). This issue can be addressed by carrying out *OR* computations on the natural logarithmic scale. We computed log odds ratios (*LOR*s) using *Comprehensive Meta-Analysis* software. The *LOR*s were then weighted and meta-analytically summarized using robust variance estimation (Hedges et al., 2010). For all reported effects, the summary effect and 95% CIs were converted from the log scale (*LOR*) back into the original metric (*OR*).

Filler selections. Lineups typically include at least four or five individuals who serve as fillers. The eyewitnesses should not have been previously exposed to these fillers, so a filler identification is a recognition error. Filler identifications can occur on culprit-present and culprit-absent lineups; however, for reasons discussed below, we only analyzed age differences in filler identifications for culprit-present lineups. For an effect size measure, *OR*s were calculated using the same procedure that was applied for hits.

Incorrect rejects. Eyewitnesses usually have the option of reporting that the culprit is “not present” and sometimes also have the option of reporting that they are “not sure.” For culprit-present lineups, both responses are errors. Researchers did not consistently provide the not-sure option, so we treated not-sure responses as lineup rejections to keep the studies that did provide the not-sure option comparable with studies that did not provide the option. *OR*s for incorrect rejects were computed using the same procedure applied for hits.

Correct rejects. For culprit-absent lineups, the correct decision is to reject all lineup members. In an applied setting, witnesses can make two types of false identifications on culprit-absent lineups: identification of the person under investigation (i.e., an innocent suspect) or identification of a filler. In experimental studies, eyewitness researchers often designate one of the culprit-absent lineup members to be the innocent suspect; however, this was rarely done in the studies that met our inclusion criteria. Most researchers only reported a correct rejection rate and an overall false positive rate. Given that the overall false positive rate is simply the inverse of the correct rejection rate, we only performed meta-analyses on the correct rejection rates to avoid redundancy. *OR* was the effect size measure for correct rejects.

Choosing. Choosing occurs when one of the lineup members is selected. For the present purposes, choosing rates did not take the accuracy of the choice into account (e.g., the selected lineup member may have been the culprit or may have been a filler). To assess age-related changes in choosing, we collapsed across culprit-present and culprit-absent lineups and calculated the overall proportion of choosers. Only studies that included both culprit-present and culprit-absent lineups were included in these analyses. Additional studies were excluded because the types of errors on culprit-present lineups were not reported and choosing could not be computed. *OR* was the effect size measure for choosing.

Diagnosticity. A properly constructed lineup should test the guilt of one person, the suspect (Wells & Turtle, 1986). The other lineup members should be fillers whose innocence is known. A filler identification or a rejection of all lineup members should be an indication of the suspect's innocence. Naturally, police will interpret a suspect identification as an indication of the suspect's guilt. If the identified suspect is indeed guilty (i.e., the culprit), the eyewitness has made a hit. However, if an identified suspect happens to be innocent, the eyewitness has made a false identification and the innocent suspect is at risk of wrongful conviction.

In the eyewitness identification literature, researchers commonly compute diagnosticity ratios to assess the relative likelihood that a suspect identification will be indicative of guilt. Diagnosticity ratios can be calculated by dividing the hit rate by the false identification rate (i.e., the innocent suspect selection rate). The diagnosticity ratio is conceptually identical to a widely used statistic in the medical literature that is referred to as a relative risk (*RR*; Tredoux, 1998). We calculated the *RR* statistic to assess age differences in diagnosticity because it has known sampling distributions and established methods of computing confidence intervals.

The *RR* statistic represents the ratio of two risks. Applied to eyewitness identification data, a risk can conceptualized as the likelihood of a suspect identification and can be calculated by dividing the number of suspect identifications by the total number of identifications. The guilty suspect identification rate (i.e., the hit rate) would be the risk for culprit-present lineups and the innocent suspect misidentification rate (i.e., the false alarm rate) would be the risk for culprit-absent lineups. An innocent suspect was not designated in most of the studies we summarized, so we calculated false alarm rates by dividing the false positive rate for all lineup members in the culprit-absent condition by the number of lineup members. We calculated *RR*s such that the value represented the ratio of hits to false alarms. For example, an *RR* of 5.00 would indicate that the hit rate was five times greater than the false alarm rate.

Age differences in *RR*s were meta-analyzed as follows: For each study that included culprit-present and culprit-absent conditions, we calculated *RR*s and 95% CIs for each age group (and for each condition) using Comprehensive Meta-Analysis software. Then, for all age comparisons, a statistic known as the ratio of relative risks (*RRR*; Altman & Bland, 2003) was calculated as an effect size for the difference between the *RR*s for two age groups. The logarithmic transformation of the *RRR*s was then meta-analyzed using robust variance estimation (Hedges et al., 2010). The *RRR*s were then converted from log scale back into their original metric.

Sensitivity (*d'*). We calculated sensitivity (discriminability) for all studies that included both culprit-present and culprit-absent conditions using the formula *d'* = *z*H - *z*FA (MacMillan & Creelman, 1991; Mickes et al., 2014), where *z*H refers to the z transformed hit rate and *z*FA refers to the *z* transformed false alarm rate. The vast majority of the studies that met our inclusion criteria tested memory for a single target person, which required sensitivity to be calculated at the group level. Following Clark (2012), we treated the proportion of correct identifications in the culprit-present condition as the hit rate and the proportion of innocent suspect selections in the culprit-absent condition as the false alarm rate. Consistent with the diagnosticity analyses, the false alarm rate was estimated by dividing the false positive rate in the culprit-absent condition by the number of lineup members.

Given that only a single *d'* score is available for each group, inferential statistics that would normally be used to compare sensitivity across conditions cannot be readily applied. However, Gourevitch and Galanter (1967) describe a method that can be used to estimate the variance of a single *d'* score, which in turn allows for calculation of inferential statistics between two group *d'* scores. In addition, obtaining the variance of group *d'* scores enables calculation of the effects size measure Hedges' g, which is Cohen's d with a correction applied to eliminate a slight bias that is characteristic of the uncorrected Cohen's d (Hedges, 1981). Hedges' g scores were computed as an effect size for the difference between group *d'* scores, which were then weighted and summarized using robust variance estimation (Hedges et al., 2010).

When calculating *d'*, hit or false alarm rates of 0 or 1 are problematic because $z(0)$ and $z(1)$ are undefined. Accordingly, when a rate of 0 or 1 was extracted from a primary study, a previously recommended correction was applied (Snodgrass & Corwin, 1988). Specifically, 0.5 was added to response frequency (e.g., the number of hits) and 1.0 was added to the group frequency (i.e., the number of participants). Whenever the correction was applied for one age group, it was also applied for the comparison age group. For example, Lindsay et al. (1995) observed a hit rate of 0 for adults in the sequential condition and a hit rate that was greater than 0 for children in the sequential condition. Although *d'* could be calculated for the child group without a correction, we applied the correction to both groups to keep the comparison of corrected rates as equivalent as possible to the comparison of uncorrected rates.

Suspect bias ($c_{suspect}$). In applications of signal detection theory, sensitivity analyses are typically accompanied by a measure of response bias. In a typical yes/no paradigm, a response bias measure indicates the participant's general inclination to respond "yes" or "no." β and *c* are two widely used response bias metrics, but *c* is generally preferable because (unlike β) it is unaffected by changes in *d'* (Stanislaw & Todorov, 1999). We calculated a measure of response bias using the formula provided by MacMillan and Creelman (1991): $c = -0.5(z\text{H} + z\text{FA})$. Consistentwith the sensitivity calculations, zH refers to the z transformed correct identification rate for the culprit-present condition and zFA refers to the z transformed innocent suspect misidentification rate for the culprit-absent condition.

Also consistent with the sensitivity calculations, response bias was calculated at the group level. To produce variance estimates for group-level *c* scores, MacMillan and Creelman

(1990; cf., Banks, 1970) note that the Gourevitch and Galanter (1967) approximation of variance for *d'* can be used in the formula var(*c*) = var(*d'*)/4. These variance approximations were used to produce the Hedges' g values for age differences in c scores, which were weighted and summarized using robust variance estimation (Hedges et al., 2010).

An important distinction between response bias for old/new recognition data and response bias for lineup data needs to be emphasized. On traditional old/new recognition tasks, c scores less than zero indicate bias toward responding "old" and c scores greater than zero indicate bias toward responding "new." Therefore, response bias for old/new tasks can be interpreted as an indication of the general inclination to choose. In the present meta-analysis, the general inclination to choose a lineup member can be inferred from the previously described outcome that was referred to as "choosing." However, the response bias measure we calculated should not be interpreted as an indication of a general inclination to choose any one of the lineup members. On the contrary, it should be interpreted as the more specific inclination to choose the suspect. Accordingly, a *c* score less than 0 would indicate a bias toward identifying the suspect and a c score greater than 0 would indicate a bias toward not identifying the suspect. To emphasize how response bias for lineup data should be interpreted, it is henceforth referred to as suspect bias (c_{suspect}).

Moderator Analyses

Metaregression is available for conducting moderator analyses using robust variance estimation. Metaregression is commonly performed with continuous moderator variables and can also be performed with categorical moderator variables via dummy coding. When performed with odds ratio as the outcome variable, metaregression produces a coefficient that can be interpreted as a ratio of odds ratios (Higgins & Green, 2011). For a categorical moderator variable such as publication status, the regression coefficient would represent the ratio of the odds ratio for age differences in published studies and the odds ratio for age differences in unpublished studies. For a continuous moderator variable such as publication year, the regression coefficient would represent the change in age differences in relation to each increase in year. We used metaregression to examine four covariates, which are described below.

Publication status. All studies were coded according to whether they had been published or not. Data obtained from journal articles or book chapters were coded as published. Data from unpublished theses, dissertations, or manuscripts were coded as unpublished.

Publication year. All studies were coded for the year in which they were published, if at all. Publication year was treated as a continuous variable. The year for unpublished studies that were in preparation or under review for publication was coded as 2015. However, if a thesis was unpublished for more than 2 years, the year of thesis completion was recorded as the publication year.

Lineup procedure. The lineup procedure was coded into the four categories: (a) simultaneous, (b) sequential, (c) elimination, and (d) showup. Different procedures are theorized to encourage different decision strategies. Wells (1984) distinguished between relative and absolute judgment strategies. Relative judgments involve comparing the lineup members with one another and deciding which one looks most like the culprit. Absolute judgments involve comparing each lineup member to a representation of the culprit in memory. The simultaneous procedure involves presentation of multiple lineup members at once, which is theorized to encourage (or at least allow for) relative judgments (Wells, 1984). The sequential procedure involves presentation of multiple lineup members one after the other, which is theorized to encourage absolute judgments (Lindsay & Wells, 1985). The elimination procedure involves two stages that are designed to encourage first relative, then absolute judgments (Pozzulo & Lindsay, 1999). First, witnesses view all the lineup members simultaneously and select the lineup member who most resembles the culprit (relative judgment). Next, all the nonchosen lineup members are removed and the witness makes a yes/no judgment about whether or not the most similar lineup member is indeed the culprit (absolute judgment). The second stage of the elimination lineup is similar to a fourth procedure, the showup test. Showups, which involve presentation of a single person and asking if he or she is the culprit, are theorized to encourage absolute judgments (Steblay, Dysart, Fulero, & Lindsay, 2003).

Young adult comparison group. The literature search revealed only one direct comparison between children and older adults. However, we used metaregression to examine whether any age differences between young adults and children were reliably different than age differences between young adults and older adults. This analysis allowed for exploration of how the child and older adult groups compared to groups that were sampled from the same population (young adults).

Results

The analyses are organized into main effects for comparisons between young adults and children/older adults, main effects for comparisons between younger and older children, and moderator analyses. Unless otherwise indicated, the analyses were performed using robust variance estimation (Hedges et al., 2010). All references to specific age ranges for groups refer to the range of mean ages. If mean age was not reported in the primary study, the median of the range of ages was used. For all null hypothesis significance tests, t values are reported as absolute values and alpha was set at .05. *OR*s or *RR*s below 1 were converted to their inverse to facilitate comparison with *OR*s or *RR*s above 1. All effect sizes reported in text are accompanied by 95% confidence intervals [*LL*, *UL*].

Main Effects of Age: Young Adults vs. Children/Older Adults

This section reports age differences between young adults and children/older adults. For each outcome, Table 1 presents the number of outlying effect sizes removed (Outliers), the number of studies (*m*) and the number of effect sizes after removing outliers (*k*), the weighted means for the two groups compared, the effect size (ES) and 95% confidence intervals (*LL*, *UL*), the significance test (*t*, *df*, *p*), and the heterogeneity indices (τ^2, I^2). For each outcome, young adults were first compared with all child comparison groups (M_{age} = 4-17 years) and all older adult groups (M_{age} = 45-77 years) and then compared with more narrowly defined groups of children (M_{age} = 5-8 and 9-13 years) and older adults (M_{age} = 68-77 years)[1].

Lineup response outcomes. Hit rate analyses indicated young adults were more likely than their older and younger counterparts to correctly identify the culprit. The odds of a hit were 1.42 [1.20, 1.69] times greater for young adults relative to children aged 4-17 and 1.71 [1.37, 2.14] times greater for young adults relative to older adults aged 45-77. In the comparisons between young adults and child groups with more narrow age ranges, the advantage in hits for young adults tended to decrease as the age of the child group increased. The odds of a hit for young adults were 1.51 [1.04, 2.18] times greater than the odds for children aged 5-8[2], whereas

[1] Comparisons between young adults and three additional age groups were attempted (M_{age} = 3-4, 14-17, and 45-48 years); however, these analyses are not reported because too few studies were available to compute trustworthy significance tests (i.e., df < 4; Tipton, 2014). Descriptive statistics for comparisons with these three groups are reported in Table 1.

[2] Relative to when outliers were excluded, the advantage in hits for young adults relative to children aged 5-8 was larger, but also had more variability and did not reach significance, *OR* = 1.70 [0.97, 2.97], $t(15) = 2.01$, $p = .06$.

the odds for young adults were 1.22 [1.01, 1.47] times greater than the odds for children aged 9-13. The largest difference in hits was in the comparison between young adults and older adults aged 68-77, *OR* = 1.95 [1.54, 2.48].

When presented with a culprit-present lineup, a witness could err by selecting one of the innocent lineup members (filler selection) or by selecting none of the lineup members (incorrect reject). Both children and older adults were consistently more likely than young adults to select a filler. Relative to young adults, the odds of a filler selection were 1.72 [1.34, 2.20] times greater for children aged 4-17 and 2.37 [1.68, 3.35] times greater for older adults aged 45-77. Similar effects were observed in the comparisons between young adults and child/older adult groups with more narrow age ranges. Children aged 4-17 were also more likely than young adults to incorrectly reject the lineup, *OR* = 1.26 [1.03, 1.54]; however, the effect was not significant for the comparisons involving children aged 5-8 and 9-13[3]. Incorrect reject rates were also not significantly different in any of the comparisons between young adults and older adults.

When presented with a culprit-absent lineup, the correct decision is to reject all of the lineup members. On the correct reject outcome, the odds were always significantly greater for young adults relative to children or older adults. The odds of a correct reject for young adults were more than double the odds for children aged 4-17, *OR* = 2.20 [1.55, 2.66]. The advantage for young adults was larger in the comparison with children aged 5-8, *OR* = 2.75 [1.70, 4.44], than in the comparison with children aged 9-13, *OR* = 2.04 [1.44, 2.89]; however, both effect sizes were substantial. The odds of a correct reject for young adults were 2.03 [1.52, 2.71] times greater than for older adults aged 45-77. The effect was marginally larger when young adults were compared with older adults aged 68-77, *OR* = 2.14 [1.58, 2.86].

[3] The difference in incorrect rejects between young adults and children aged 4-17 was also not significant when outliers were included, *OR* = 1.33 [0.97, 1.83], $t(17) = 1.91$, $p = .07$.

Choosing represents the overall rate at which lineup members were selected, collapsed across culprit-present and culprit-absent lineups. The analyses for this outcome indicated the odds of choosing were significantly greater for children aged 4-17 and older adults aged 45-77 relative to young adults, *OR* = 1.38 [1.20, 1.61] and *OR* = 1.56 [1.23, 1.98], respectively. The effect size was larger for the comparison between young adults and children aged 5-8, *OR* = 1.72 [0.88, 3.33], than for the comparison between young adults and children aged 9-13 years, *OR* = 1.39 [1.11, 1.75]; however, there were fewer effect sizes to summarize in the comparison involving the 5-8 years group and only the comparison involving the 9-13 years group was significant. The effect for the comparison between young adults and older adults aged 68-77 was also significant, *OR* = 1.69 [1.27, 2.27].

The response rate analyses consistently indicated that young adults perform better than their younger and older counterparts on eyewitness identification tasks. Relative to children and older adults, young adults were more likely to identify the culprit and less likely to select a filler on culprit-present lineups. In addition, young adults were consistently more likely than children and older adults to reject culprit-absent lineups. Analysis of the choosing outcome indicated that children and older adults selected lineup members at a higher rate than did young adults, which would explain the large age differences in correct rejections and smaller age differences in correct identifications.

Sensitivity, diagnosticity, and suspect bias. As can be seen in Table 1, the sensitivity analyses indicated that young adults were better able than children (*d’* = 1.69 vs. 1.31) and older adults (*d’* = 1.54 vs. 0.95) to discriminate between guilty and innocent suspects. In each of the meta-analytic comparisons, *d’* scores for young adults were significantly greater than *d’* scores for children and older adults; however, all Hedges' *g* values were between 0.11 and 0.19, which suggests a small but reliable advantage in discriminability for young adults relative to children and older adults[4]. The diagnosticity analyses led to conclusions similar to those for the sensitivity analyses, with young adults consistently producing larger diagnosticity ratios relative to both children and older adults.

Young adults, children, and older adults were all biased towards not identifying the suspect (i.e., all suspect bias values were greater than zero). The positive values for suspect bias are not particularly surprising given that the chance likelihood of a suspect identification is usually 17% for any witness who chooses one of the lineup members and witnesses also typically have the option of not choosing any of the lineup members. Therefore, negative values on this measure should not be expected. Although all groups were biased towards not identifying the suspect, the bias was significantly greater for older adults than for young adults. By contrast, no significant differences in suspect bias were observed in any of the comparisons between young adults and children.

Main Effects of Age: Younger Children vs. Older Children

The next set of meta-analyses summarized all effect sizes for comparisons between two child groups of different ages. The dataset for these analyses comprised any effect size for a comparison between two child groups that differed in mean age. The groups compared are referred to as "younger" and "older". To be included in these analyses, the mean age of the older group needed to be (a) greater than the mean age of the younger group and (b) less than 18 years.

[4] All differences in sensitivity were significant when outliers were excluded. When outliers were included, the advantage in sensitivity for young adults relative to children aged 5-8 was the only difference that did not reach significance, $g = 0.16$ [-0.01, 0.33], $t(7) = 2.28$, $p = .06$.

Table 2 presents descriptive and inferential statistics for comparisons between child groups of different ages (younger vs. older). Reliable age differences were found for responses on culprit-present lineups and for sensitivity and diagnosticity. The odds of a hit were significantly greater for older children than for younger children, *OR* = 1.46 [1.22, 1.75]. Younger children made both types of culprit-present lineup errors more often than did older children, with a larger effect size for filler selections than for incorrect rejects, *OR* = 1.64 [1.32, 2.00] and *OR* = 1.27 [1.06, 1.54][5], respectively. The odds of a correct reject were greater for older children than for younger children, but the effect was not significant, *OR* = 1.23 [0.97, 1.57]. Choosing rates for older and young children also did not significantly differ, *OR* = 1.06 [0.90, 1.26]. Sensitivity and diagnosticity values were both significantly greater for older children relative to younger children, *g* = 0.05 [0.01, 0.11] and *RRR* = 1.22 [1.00, 1.48], respectively. Both younger and older children had conservative suspect biases that were not significantly different, *g* = -0.03 [-0.08, 0.01], indicating both groups exhibited a similar tendency to not identify the suspect. The age-related increase in sensitivity and diagnosticity suggest a developmental improvement in identification performance.

Mean Age and Mean Age Difference

Meta-regression can be applied using robust variance estimation to assess the impact of covariates. Moreover, robust variance estimation allows a covariate's impact to be parsed into between-study effects and within-study effects (Uttal et al., 2013). For example, in some studies adults were compared with 5-year-olds and 10-year-olds, and in other studies adults were

[5] Including outliers produced a nonsignificant difference in incorrect rejects, *OR* = 1.15 [0.92, 1.43], $t(23) = 1.26$, $p = .22$.

compared with only 5-year-olds or only 10-year-olds. If the studies that compared adults to both child groups produce larger effects for adults compared to 5-year-olds than for adults compared to 10-year-olds, this would be considered a within-study effect. Conversely, if the studies that only compared adults to 5-year-olds yield larger effects than studies that only compared adults to 10-year-olds, this would be considered a between-study effect.

We separated between- and within-study effects for two covariates. First, we treated the child group's mean age as a covariate in the comparisons between young adults and children aged 4-17. Second, for the comparisons between younger and older child groups, we treated the mean difference in age between groups as a covariate. For both sets of analyses, not all outcomes had enough comparisons between young adults and children to perform the meta-regression separately for between- and within-study effects. Specifically, for all outcomes that required data for culprit-present and culprit-absent conditions to create a composite measure (i.e., choosing, sensitivity, diagnosticity, and suspect bias), separating the between- and within-study effects caused the degrees of freedom to be less than 4, which tends to inflate the Type 1 error rate (Tipton, 2014). This was also the case for the correct reject outcome in the first set of analyses. Accordingly, between- and within-study effects were combined for outcomes with an insufficient number of studies for separating the two types of effects.

Table 3 presents the meta-regression coefficients, 95% CIs, and significance tests for the two covariates. For the mean age covariate, a significant association was observed for hits. The age difference in hits between young adults and children was negatively associated with the age of the child comparison group, indicating the difference in hits between young adults and children tended to decrease as the age of the child group increased. This effect was only significant when between- and within-study effects were combined; however, the within- and between-study associations were in the same direction and had similar magnitudes, suggesting a similar trend for both types of effects. The size of the age differences was not significantly associated with the child group's mean age for any of the other outcomes.

For the mean age difference covariate, significant associations were observed for hits and diagnosticity. The size of the age difference in hits was positively associated with the difference in mean age between the younger and older groups; however, the association was only significant for the within-study effects. The magnitude of the association for within-study effects was also larger than for between-study effects. This suggests the difference in hits between older children and younger children tended to increase as the age difference between the two age groups increased, but this effect was only reliable when the experimental procedure was controlled. When between-study confounds were present, the association was not reliable. For diagnosticity, the size of the age difference was significantly associated with the mean age difference. The number of studies was insufficient for separating between- and within-study effects for diagnosticity.

Publication Year and Publication Status

The influence of publication year and publication status on the size of age differences was assessed for two datasets, one containing comparisons between young adults and both children (M_{age} = 4-17) and older adults (M_{age} = 45-77) and the other containing comparisons between younger and older children. Meta-regression was performed to simultaneously assess the impact of publication year and publication status. Entering two covariates into the meta-regression together ensures that any associations observed for one covariate have been controlled for the influence of the other covariate. The young adult/child comparisons and the young adult/older adult comparisons were combined because the effects for these groups were similar. In addition, performing these analyses on the full dataset, rather than datasets parsed into smaller components for more focused analyses, has the practical benefit of ensuring that all analyses had at least 4 degrees of freedom, which is the minimum required to obtain trustworthy significance tests using robust variance estimation (Tipton, 2014).

Table 4 presents the meta-regression coefficients, 95% *CI*s, and significance tests for publication year and publication status. Publication year was significantly associated with age differences in correct rejects. The association was significant for comparisons between young adults and children/older adults, ratio of *OR*s = 0.98 [0.95, 1.00], $t(7) = 2.63$, $p = .032$, and for comparisons between older children and younger children, ratio of *OR*s = 0.96 [0.93, 0.98], $t(12) = 3.42$, $p = .005$. We explored the association for the first dataset by computing the difference in correct rejects (young adult rate minus child/older adult rate) for each effect size and plotted the difference scores according to publication year. Figure 1 shows a negative association in which the young adult advantage in correct rejects decreases as the year of publication increases. A similar association was observed in the dataset comparing older children and younger children, such that the advantage in correct rejects for older children decreases as the year of publication increases. Thus, for both datasets, the age difference in correct rejects was larger in older studies than in newer studies.

The only additional significant association for publication year was for choosing rates in the young adult versus children/older adult dataset, ratio of *OR*s = 1.02 [1.01, 1.04], $t(5) = 4.12$, $p = .01$. Contrary to correct rejects, which were more likely for young adults, choosing was more likely for children/older adults. Accordingly, we interpreted this association by subtracting the young adult choosing rate from the child/older adult choosing rate and plotting the difference scores according to publication year (Figure 2). The plot shows a negative association in which age differences in choosing tend to be larger in older studies than in newer studies. Thus, the finding that children/older adults choose more (and correctly reject less) has been diminishing with time.

The meta-regressions on the publication status covariate yielded only one significant association. For the diagnosticity outcome in comparisons between young adults and children/older adults, the advantage for young adults was larger in published studies (*RRR* = 1.92 [1.54, 2.40], *m* = 30, *k* = 92) than in unpublished studies (*RRR* = 1.28 [0.95, 1.72], *m* = 9, *k* = 21) which yielded a ratio of *RRR*s that was significantly different from unity, 1.54 [1.07, 2.20], $t(11) = 2.67$, $p = .02$. Although this was the only significant association, the associations between publication status and age differences on three additional outcomes approached significance. Specifically, age differences were marginally larger in published relative to unpublished studies in the analyses of sensitivity (published: *g* = 0.15 [0.12, 0.19], *m* = 30, *k* = 89; unpublished: *g* = 0.08 [-0.01, 0.18], *m* = 9, *k* = 21), filler selections (published: *OR* = 2.19 [1.69, 2.86], *m* = 35, *k* = 121; unpublished: *OR* = 1.53 [1.01, 2.31], *m* = 8, *k* = 21), and correct rejects (published: *OR* = 2.28 [1.81, 2.86], *m* = 37, *k* = 98; unpublished: *OR* = 1.41 [1.17, 1.70], *m* = 8, *k* = 17). None of the associations between publication status and age differences in the younger versus older children dataset were significant.

Lineup procedure

Moderator analyses were performed to assess for procedural influences on age differences between young adults and children/older adults. Four identification procedures were examined: simultaneous, sequential, elimination, and showup. The simultaneous, sequential, and elimination procedures all involve presenting multiple lineup members for identification. The showup procedure involves presenting a single suspect for identification. Accordingly, moderator effects of lineup procedure were organized as comparisons between the three types of lineups and comparisons between lineups and showups.

Two sets of analyses were conducted to assess for differences between the three lineup procedures, one for between-studies effects and one for within-study effects. For the analyses of between-study effects, the lineup procedures were dummy coded and meta-regressions were performed on eight outcomes (hits, filler selections, incorrect rejects, correct rejects, choosing, sensitivity, diagnosticity, and suspect bias). For the comparison between young adults and children, the simultaneous procedure was designated as the procedure to be compared against the elimination and sequential procedures because the latter two procedures were rarely compared with each other. For the comparison between young adults and older adults, only the simultaneous and sequential procedures were compared because older adults were only tested with elimination lineups in one study. Regardless of which groups were compared or which outcome was tested, none of the meta-regressions indicated that lineup procedure was a significant moderator of age differences. Given that meta-regression is a commonly under-powered statistical technique (Hedges & Pigott, 2004), the absence of significant differences does not necessarily rule out lineup presentation as a moderator of age differences. Nevertheless, these between-study analyses provided no indication that lineup presentation influences age differences in eyewitness identification.

For analyses of within-study lineup procedure effects, only studies that directly compared two or more lineup procedures were included. There were too few effect sizes to perform meta-regression with robust variance estimation, so these analyses were performed using categorical moderator analysis in *Comprehensive Meta-Analysis* software. Contrary to the robust variance estimation technique, this more conventional approach does not have assumptions regarding a minimum number of degrees of freedom. However, some adjustments to the dataset were required because the current version of *Comprehensive Meta-Analysis* software does not provide any specific options to account for effect size dependence. In particular, to address effect size dependence in studies with comparisons between adults and more than one child comparison group (e.g., 5-year-olds and 10-year-olds), the effect sizes were recalculated with the child groups collapsed. Some effect size dependence was still present in studies employing factorial designs with lineup procedure and another independent variable, but we did not collapse across the second independent variable to avoid losing the information it provided. Weighted means and effect sizes were computed using the fixed-effect model to ensure that study weights were not directly linked to the study's number of effect sizes. The moderator analyses were performed using a mixed-effects model. We limit our discussion of procedural influences on age differences to comparisons between young adults and children because within-study comparisons between young adults and older adults were particularly rare. For those interested, the comparisons between young adults and older adults are reported in Appendix B.

Table 5 presents the within-study effects of lineup procedure on age differences between young adults and children. On the whole, the moderator tests suggested that the identification procedure rarely had a reliable influence on age differences. However, these analyses typically had a limited number of comparisons, so the absence of a significant moderator test should not be interpreted as conclusive evidence that lineup procedures do not affect some age groups differently than others. Moreover, as will become clear in our discussion of the results, some of the procedures produced age differences that were in clear contrast to those reported in the main effect analyses.

Lineups vs. showups. In the comparison between lineups and showups, the identification procedure only had a significant moderator effect on age differences in suspect bias, $Q(1) = 4.18$, $p = 0.04$. The descriptive information in Table 5 shows that on lineups, adults ($c_{suspect} = 0.82$) were more biased towards not identifying the suspect than were children ($c_{suspect} = 0.25$). On showups, the difference in suspect bias was even larger, as children were biased towards identifying the suspect ($c_{suspect} = -0.18$) and adults were strongly biased away from identifying the suspect ($c_{suspect} = 0.88$). Of all the suspect bias analyses, children's showup identifications represent the only case in which witnesses were biased toward identifying the suspect.

The differences in suspect bias between children and adults in showups relative to lineups draws attention to some particularly atypical identification responses in the small number of studies ($m = 3$) that have employed showups with child witnesses. Young adults were nearly unanimous (94%) in their decision to correctly reject a culprit-absent showup, whereas less than two thirds of those in the child comparison groups (62%) were inclined to do the same. A 32% difference in correct rejects is much larger than the difference between children and adults in the full dataset of lineup studies (14%, see Table 1); however, in the showup studies, very large age differences in correct rejects were also observed in the lineup conditions (young adults = 70%; children = 40%). The culprit-present conditions also produced somewhat peculiar results. On culprit-present showups, the odds of a hit was 3.55 [1.70, 7.45] times greater for children (77%) than for adults (49%). Similarly, in this subset of studies, children were significantly more likely than young adults to make a hit from a culprit-present lineup (64% vs. 53%, respectively), *OR* =

1.64 [1.01, 2.68]. Thus, although children were clearly biased towards identifying the suspect in the showup conditions, children in the lineup comparison conditions also seemed to have selected the suspect more than usual.

Simultaneous vs. elimination[6]. In the comparison between simultaneous and elimination procedures, none of the moderator tests were significant. The elimination procedure was designed to improve children's correct rejection rates, but these analyses showed that adults had reliably higher correct reject rates for both the simultaneous procedure, *OR* = 2.17 [1.33, 3.54], and the elimination procedure, *OR* = 2.92 [1.72, 4.95]. The only benefit of the elimination procedure for children seemed to be in hit rates, which were comparable for children (48%) and young adults (51%), *OR* = 1.14 [0.71, 1.83]. However, the comparable hit rates seem to be a consequence of the elimination lineup producing a higher rate of choosing in children (54%) relative to young adults (38%), *OR* = 1.96 [1.42, 2.71]. Analyses of sensitivity indicated the advantage for adults over children was virtually identical across the simultaneous (d' = 1.98 vs. d' = 1.49; g = 0.14 [-0.02, 0.29]) and elimination (d' = 1.79 vs. d' = 1.31; g = 0.13 [-0.03, 0.28]) procedures.

Simultaneous vs. sequential. In the comparison between simultaneous and sequential procedures, the analysis of correct rejects yielded the only significant moderator effect. This effect was indicative of a larger advantage for young adults relative to children in correct rejects for lineups presented sequentially (73% vs. 33%, respectively) than for lineups presented simultaneously (55% vs. 40%, respectively), $Q(1) = 6.36$, $p = .01$. There was also a large difference in sensitivity for sequential presentation (young adult: d' = 1.73; children: d' = 1.03),

[6] Pozzulo and Lindsay (1999) initially proposed two types of elimination procedures: fast and slow. The slow elimination procedure has not been employed since, so only comparisons between simultaneous and fast elimination lineups were included in these analyses.

compared with only a small difference in sensitivity for simultaneous presentation (young adult: *d'* = 1.64; children: *d'* = 1.58); however, the moderator test was not significant, $Q(1) = 2.48$, $p = .12$. Hit rates for young adults and children on simultaneous lineups were also comparable in this subset of the data (59% vs 60%, respectively).

The finding of comparable *d'* scores for children and young adults on simultaneous lineups in the simultaneous-sequential comparisons is inconsistent with the main effect analyses, which consistently showed better discriminability for young adults. Given that the main effect analyses included data for all three lineup procedures, it is possible that children perform as well as adults on simultaneous lineups and the difference between children and adults in the main effects was driven by age comparisons on other procedures. To assess this possibility, we examined the difference between young adults and children for the full set of studies that used simultaneous presentation. These analyses, which were performed using robust variance estimation, show that simultaneous presentation led to an advantage for young adults over children in hits, correct rejects, sensitivity, and diagnosticity (Table 6). Thus, the lack of the age differences for simultaneous lineups in the subset of simultaneous-sequential comparisons was not representative of age differences in performance on simultaneous lineups more generally.

Differences between Young Adults' Comparison Groups (Children vs. Older Adults)

The main effect analyses indicated that young adults performed better than both children and older adults, which begs the question: Do children and older adults differ from one another? Ideally, this question would be answered through a meta-analytic summary of direct comparisons between these two groups. Unfortunately, our literature search revealed only one direct comparison between children and older adults on an eyewitness identification task (Morten, 2014), which precluded a direct meta-analytic comparison. Although direct comparisons between children and older adults were rare, our dataset contained numerous comparisons between young adults and children or older adults. Therefore, moderator analyses can be performed to examine whether the advantage for young adults is reliably larger when compared to either children or older adults.

To perform meta-regression with robust variance estimation, dummy codes were assigned to distinguish between comparisons between young adults and children (0) and comparisons between young adults and older adults (1). The meta-regression coefficients in Table 7 revealed one significant difference. Specifically, the difference in *d'* scores for comparisons between young adults ($M = 1.54$) and older adults ($M = 0.95$) was significantly larger than for comparisons between young adults ($M = 1.69$) and children ($M = 1.31$), $t(23) = 2.30$, $p = .03$. Although this analysis is not sufficient to infer greater discriminability in children relative to older adults, it shows that decrements in discriminability relative to young adults are reliably larger in comparisons with older adults than in comparisons with children.

Discussion

This meta-analysis of eyewitness identification across the life span revealed clear age differences in identification responses. Relative to both children and older adults, young adults evinced a higher correct identification rate on culprit-present lineups and a higher correct rejection rate on culprit-absent lineups. Signal detection analyses allowed us to explore the reasons for age differences in identification performance. As anticipated, both older adults and children chose from lineups at a significantly higher rate than did young adults. However, the ability to discriminate between innocent and guilty suspects was significantly greater for young adults than for older adults and children. These findings provide clear mechanistic evidence for the observed age differences in accuracy: Young adults' superior performance is a result of an enhanced ability to discriminate between guilty and innocent suspects, not just a more conservative selection criterion.

The advantage for young adults was remarkably consistent across comparisons with different age groups. Relative to young adults, both children and older adults were less likely to correctly identify the culprit and more likely to identify a filler from culprit-present lineups, and both groups were also significantly less likely to correctly reject a culprit-absent lineup. When we compared response patterns across children of different ages to explore developmental differences, the results were equally clear: Lineup identification accuracy improved throughout childhood. The size of the difference in hit rates observed between young adults and children decreased as the age of the child group increased. Similarly, the size of the difference in correct reject rates between young adults and children decreased as children's age increased. Further, the size of the difference in hits between older and younger children increased as age differences increased. In totality, the data converge on the conclusion that children's discriminability increases throughout childhood until it begins to resemble that of young adults.

Though intuitive, some of these findings stand in contrast to the previous suggestions in the literature. The results from the only meta-analytic data available on the topic (Pozzulo & Lindsay, 1998) indicated that children "can" identify the culprit as effectively as young adults, provided that the culprit is in the lineup. Despite children's overall higher level of choosing, the present analyses make it clear that this choosing mostly led to filler selections. Contrary to the existing consensus in the literature, we found higher correct identification rates for adults than for children.

Implications for Life Span Theory Development

This meta-analysis provides the first review of eyewitness identification across the life span. The findings related to children and older adults when compared with young adults are clear, and we draw conclusions with confidence. Crucially, however, direct meta-analytic comparisons between children and older adults were not possible due to a lack of direct comparisons. Thus, comparisons of the similarity of these two populations must come from theoretical predictions and behavioral observations relative to the performance of young adults. As outlined above, there were many similarities in children's and older adults' lineup performance. These similarities indicate that identification decisions may be affected by similar processes in cognitive development and cognitive aging. However, there are both theoretical and observational reasons to posit that children and older adults' identification decisions may be affected by different processes.

Memory processes. We previously argued that older adults and children may be more likely than younger adults to rely on familiarity processes, and that older adults may also be further disadvantaged due to weaker associative processes (Brainerd & Reyna, 2004; Shing et al., 2010, 2008). In a clear example of how such dissociation may play out, Shing, Werkle-Bergner, Li, and Lindenberger (2008) manipulated strategy instructions and associative demand in a recognition memory experiment. Children's false alarms were reduced with strategy instruction, implying that the memory trace, or associative binding, was not the primary source of weakness (see also Fandakova et al., 2014). In contrast, older adults did not benefit as much from strategy instruction and their performance was especially weak when associative demands were strong. Shing et al. (2008) hypothesized that older adults' high false alarm rate may have been a result of less distinct memory traces, resulting from a deficiency in the associative component combined with diminished strategic control at retrieval. This pattern indicates that children's memory traces, with assistance during retrieval, may be more comparable with young adults'. These indistinct memory traces observed in older adults that result from more limited associative processes (Shing et al., 2008) may contribute to relatively greater difficulty in discriminability.

Although a meta-analytic comparison was not possible, the analyses of discriminability and suspect bias provide some evidence for potential differences between children and older adults. The signal detection analyses showed that older adults' discriminability was more different from young adults' than was children's, suggesting that different processes may contribute in variable ways to identification decisions made by each of these populations. Borne out, this observation would contribute to a larger body of work supporting similar behaviors with different underlying processes during cognitive development and cognitive aging (see Craik & Bialystok, 2006; Sander et al., 2012; Shing et al., 2010, 2008).

Social and strategic processes. As we discussed in the introduction, a substantive focus in the child witness literature has been on social pressures, largely because the historic view of children as witnesses has been interlaced with the literature on children's suggestibility (Bruck & Ceci, 1999). Concordant evidence from the eyewitness identification literature supports the notion that children's identification decisions can be influenced by social conditions. Children show improvement in identifications with unbiased instructions (Pozzulo & Dempsey, 2006), salient lineup rejection options (Dunlevy & Cherryman, 2013; Karageorge & Zajac, 2011; Zajac & Karageorge, 2009), and a lack of overt social pressures (Lowenstein et al., 2010). Though much less research has been conducted with older adults in this area, that which has been conducted indicates that older adults seem relatively unaffected by biased instructions (Rose, Bull, & Vrij, 2005; Searcy, Bartlett, & Memon, 2000; Wilcock, Bull, & Vrij, 2005) and any effect of salient rejection options has been inconsistent (Gentle, 2012; Havard, n.d.). Given that both children and older adults tend to choose at higher rates than young adults, and unbiased instructions appear to assist children but not older adults, perceived social pressures may contribute more to children's poorer performance than to older adults' performance. This pattern speaks to children's vulnerability to the implicit demands of the lineup task and also to the importance of applying best practice in lineup procedures.

Further support for children's greater susceptibility to social pressures can be found in the meta-analytic comparisons of identification procedures. Generally, all age groups were biased against identifying the suspect, though the bias was strongest in older adults. However, when showups were considered (a procedure that many researchers consider to be inherently biased), children were biased toward identifying the suspect, whereas young adults were biased away from identifying the suspect. As with the above examples, suspect bias in older adults was not impacted by the identification procedure, thus further supporting the idea that interventions targeting strategic processes are less effective with older adults.

If children's errors are reduced with interventions targeted at social processes, why do such interventions fail to bring children to the same levels as young adults? Though differences in associative processes between children and young adults are not always evident, differences in strategic and recollective processes are clear (Brainerd & Reyna, 2004; Shing et al., 2010, 2008). Social interventions appear to reduce the magnitude of the differences by targeting children's strategic processes, but children still choose at higher rates and make more errors than do young adults. These choices, perhaps due to greater reliance on familiarity, may also be the result of a greater willingness to guess.

Manipulations of the similarity between fillers and the target provide some insight into guessing tendencies. A witness prone to high rates of guessing is likely to be more successful in a biased lineup because the likelihood that the witness will select the culprit is increased when the plausibility of fillers is decreased. In contrast, if all lineup members are plausible, guessing will be less successful because the chance of selecting the culprit is distributed across the full set of options. Thus, increasing filler similarity should have a larger effect on witnesses who are more prone to guessing (i.e., by reducing correct identifications). Consistent with the suggestion that children are more likely to guess than young adults, manipulations of filler similarity have been found to have a substantially larger effect on children than on young adults (Fitzgerald, Whiting, Therrien, & Price, 2014). Conversely, filler similarity manipulations have not been found to differentially affect older adults relative to young adults (Key et al., n.d.). Together, these findings suggest that children, but not older adults, are more prone to guess in the face of uncertainty than are young adults.

The empirical literature is consistent with the suggestion that children are influenced by social pressures to choose. Children's choosing is reduced through warnings that the culprit may be absent and increasing the saliency of the rejection option. Conversely, older adults' limited associative processes have not been clearly shown to be influenced by such manipulations. Though it is crucial to be mindful that the particular manipulations and populations in these comparisons differ, the potential differences drive home the need for strategic and systematic comparisons across the life span. There is evidence that children and older adults perform similarly, but also that there are likely differences in the underlying processes.

Lineup Procedure

One of the hopes in the child eyewitness identification literature has been that system variables could be developed to assist children in reaching adult-like performance. However, none of the lineup procedures we compared (simultaneous, sequential, showup, elimination) reduced the gap in discriminability between children and young adults.

The elimination lineup was specifically designed to accommodate children's increased propensity to choose from culprit-absent lineups. Although researchers have suggested that the elimination procedure results in accuracy rates similar to those for young adults (e.g., Pozzulo et al., 2013), the moderator analyses show that the elimination lineup does not reduce the age difference between children and young adults in discriminability. For the elimination lineup, discriminability was 1.79 for young adults and 1.31 for children, a difference of 0.48. For the simultaneous lineup, discriminability was 1.98 for adults and 1.49 for children, a difference of 0.49. These results clearly show that the elimination procedure does not improve accuracy.

None of the other procedures fared much better. The sequential procedure did not reduce the gap between children and adults in correct identifications, and the adult advantage in correct rejections increased for sequential relative to simultaneous lineups. Although the showup procedure increased children's correct identification rates more than it did for young adults', children seem inclined to make identifications from showups regardless of whether the suspect is guilty or innocent, providing further evidence of children's increased propensity to make an identification even if they are just guessing. Thus, it appears we have not yet developed a lineup procedure that improves children's discriminability, relative to young adults.

Limitations of the Meta-Analysis

Meta-analyses are often limited by the data that are available to summarize, and the meta-analysis reported here is no exception. Although we found reliable age differences in identification performance, in none of the primary studies were the same individuals tested at different stages of their lives. This universal application of cross-sectional designs limits our understanding of age-related changes in eyewitness identification. Given the resources required to conduct longitudinal research, the more pragmatic approach adopted by researchers is understandable. The benefits of following individuals from childhood to older adulthood to examine changes in eyewitness identification may not be worth the costs. However, we could only find one cross-sectional study involving children and older adults. In our search for relevant articles, we found two distinct literatures: one for younger witnesses and one for older witnesses. Given the similar identification patterns for these two groups, the reasons for this disconnect are unclear. We hope the life span approach adopted in this meta-analytic review facilitates the merging of these two literatures.

Our analyses were also limited by the nature of the eyewitness identification paradigm. In a typical eyewitness identification experiment, witnesses encounter a single target person and then complete a single identification task. The identification task typically includes at least six lineup members, one of whom may be the target. We computed signal detection measures to examine age differences in discriminability, but these measures were designed for analyzing responses to multiple trials of yes/no tasks. The method we used to calculate *d'* provides a measure of discrimination between guilty suspects from culprit-present lineups and innocent suspects from culprit-absent lineups (Mickes et al., 2014). This approach does not differentiate between the two types of errors on culprit-present lineups (i.e., filler selections and incorrect rejections) and does not include all mistaken identifications on culprit-absent lineups in the calculation of the false alarm rate. This approach corresponds with how identification responses are typically interpreted in applied settings, where only the decision for the suspect has implications for an investigation. However, alternative approaches to computing signal detection measures may better represent the decision-making strategies for eyewitness identification tasks (see Palmer & Brewer, 2012; Palmer, Brewer, & Weber, 2010).

How the data were reported in primary studies also limited our options for synthesis. From the data reported, we could extract the information necessary to compute *d'* and the diagnosticity ratio. However, when calculating these measures we did not take the witnesses' confidence into account. This is a potentially important limitation because it is possible that the inferior performance observed for a particular age group only occurs for identifications made with low confidence. In recent studies, eyewitness identification researchers have computed a diagnosticity ratio for each reported level of confidence, which can be plotted as a receiver operator characteristic (ROC) curve. In a recent report on eyewitness identification research and practices, the calculation of ROC curves was commended for providing more information than a single diagnosticity ratio (National Research Council, 2014). However, we were unable to compute ROC curves because researchers rarely reported identification rates at each level of confidence in the primary studies. To allow future meta-analyses to examine the influence of variables at different levels of confidence, researchers could report identification rates at different levels of confidence (for an example, see Wells & Penrod, 2011).

Directions for Future Research

Going forward, we encourage a life span approach to eyewitness identification research. Given the paucity of research directly comparing children and older adults, this is a critical area in need of further research. Investigating whether similar underlying mechanisms are responsible for the poor discriminability in these populations is of particular interest. Wixted and Mickes (2014) recently articulated a theoretical account of eyewitness identification that emphasizes the role that attending to diagnostic features among the lineup members can have on measures of discriminability. Thus, children's and older adults' reduced accuracy on lineup tasks may result from an increased tendency to focus on nondiagnostic features. Future research should explore this possibility.

In addition to directly comparing children and older adults, we encourage researchers to conduct more studies with underrepresented age groups. Most of the comparisons involved children (5–13 years), young adults (18 –25 years), and older adults (68 – 80 years). By contrast, research on older adolescents (15–17 years), adults in the middle years (30 –55 years), and adults in the very late years (80+ years) was lacking. Research on older adolescents may be useful for practical applications, given their relatively high likelihood of witnessing a crime. Furthermore, given that research on the effects of aging on eyewitness performance have been almost entirely limited to comparisons between one group of young adults and one group of older adults, the point at which accuracy declines occur is unclear. Additional research examining multiple age groups is needed to provide a clearer picture of when and how quickly older adults show reduced identification accuracy. Until further research has been conducted with these groups, a comprehensive understanding of eyewitness identification across the life span will remain elusive.

Research on older adult witnesses also tended to use similar methods, which limits the generalizability of the findings. In almost every study, the researchers used video events, exposure times of less than 1 min, delays of less than 24 hr, unbiased instructions, and six-member photographic lineups. The lack of live events is particularly limiting. Although video events have a role in eyewitness identification research, this is not how eyewitness encounters occur in applied settings. Clearly, there is a need for research using live events with older adults.

In addition, some potentially important moderators of age differences have received virtually no attention in the literature on age differences in eyewitness identification. For example, live lineups were not used in a single study that met our inclusion criteria. Given that live lineups are relatively common in the United States, and standard practice in other countries (e.g., Australia, South Africa), more research is needed to examine whether age differences are moderated by this variable.

The composition of the lineups is another critical variable that seems to have gone largely under the radar of researchers examining children and older adults. Many years ago, researchers called for exploration of the level of similarity among lineup members for child witnesses (Davies et al., 1989; Parker & Carranza, 1989); however, the first empirical research on filler selection procedures for children was only very recently published (Fitzgerald et al., 2014). Similarly, we found only one unpublished study examining lineup member similarity with older adults (Key et al., n.d.). Given the clear effects of similarity on identification responses that have been demonstrated with young adults (Fitzgerald, Price, Oriet, & Charman, 2013), exploring this variable with different populations is crucial.

Examining age differences in eyewitness identification for different levels of subjective confidence ratings is another important direction for future research. For young adults, increases in identification accuracy tend to coincide with increases in identification confidence (e.g., Lindsay, Read, & Sharma, 1998), provided that no post-identification feedback is given (Bradfield, Wells, & Olson, 2002). By contrast, children and older adults have been shown to have less developed metacognitive awareness on lineup tasks (Keast et al., 2007; Wylie, Bergt, Haby, Brank, & Bornstein, 2014), though research examining children's and older adults' identification confidence is sparse. The literature would benefit from additional experimental research that explores age differences at varying degrees of confidence, which would allow ratios of hit and false alarm rates to be plotted as ROC curves (see Mickes, Flowe, & Wixted, 2012).

An additional recommendation for future work is a thorough and thoughtful study space analysis (Malpass et al., 2008). This type of analysis would identify both the strengths of the existing body of work, as well as the gaps in literature. It could promote the marrying of findings from both ends of the life span through systematic identification of variables that need to be addressed. Perhaps particularly helpful for the current research, it could provide an assessment of both cross-study and cross-laboratory findings. We have attempted to highlight several areas that we perceive to be the most urgently needed directions for future work, but a study space analysis would be a welcome addition.

Practical Applications

Wells (1978) emphasized the need to understand the impact of system variables, arguing system variable research has the most potential for developing policies to reduce eyewitness errors. Nevertheless, estimator variables are typically more influential in court. When assessing identification reliability, the guidelines that have been outlined in court decisions advise consideration primarily of information processing factors that were present at the witnessed event. For example, judges often question whether the witness had a good opportunity to view the culprit. We examined another potentially useful factor to consider when determining the reliability of eyewitness identifications: the age of the witness.

Our meta-analytic review indicated that witness age is a reliable predictor of identification accuracy. Young adults were superior witnesses when compared with children and older adults. Note, however, that our analyses did not take eyewitness certainty into account. In some jurisdictions, members of the legal system are advised to take eyewitness confidence into account when assessing the reliability of lineup identifications (e.g., *Manson v. Braithwaite*, 1977). Young adults have been shown to have greater metacognitive awareness relative to children and older adults (Keast et al., 2007; Wylie et al., 2014), which may give additional reason for caution when assessing the reliability of children's and older adults' identification, even when they report high confidence.

Although our meta-analytic review suggests that identifications by young adults are more reliable than identifications by children and older adults, it also showed that identifications by witnesses of all ages can be reliable indicators of the suspect's guilt. Although clear age differences exist, the analyses showed that identifications by children and older adults have diagnostic value. Most of the diagnosticity ratios for children and older adults were around 4 –5, suggesting that a guilty suspect was 4 –5 times more likely than an innocent suspect to be identified. Thus, our meta-analytic summary of the experimental research suggests eyewitness identifications from all age groups have value.

References

References marked with an asterisk (*) indicate studies included in the meta-analysis.

Aizpurua, A., Garcia-Bajos, E., & Migueles, M. (2011). False recognition and source attribution for actions of an emotional event in older and younger adults. *Experimental Aging Research, 37*, 310–329. doi: 10.1080/0361073X.2011.568829

Altman, D. G., & Bland, M. J. (2003). Interaction revisited: The difference between two estimates. *BMJ, 326,* 219-219. doi: 10.1136/bmj.326.7382.219

Anderson, N. D., & Craik, F. I. M. (2000). Memory in the aging brain. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 411–422). New York: Oxford University Press.

*Ball, M. (2009). *The effect of stress on memory: Eyewitness performance in juveniles and young adults.* Unpublished thesis. University of Cape Town, South Africa.

Balota, D. A., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy older adults. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 395–409). New York: Oxford University Press.

Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin*, *74,* 61-99. doi: 10.1037/h0029531

Bartlett, J. C. (2014). The older eyewitness. In T. J. Perfect & D. S. Lindsay (Eds.), *The SAGE handbook of applied memory* (pp. 654–674). London, UK: SAGE.

Bartlett, J. C., & Memon, A. (2007). Eyewitness memory in young and older eyewitnesses. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Handbook of eyewitness psychology: Memory for people* (Vol. 2, pp. 309–338). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

*Beresford, J., & Blades, M. (2006). Children's identification of faces from lineups: The effects of lineup presentation and instructions on accuracy. *Journal of Applied Psychology, 91,* 1102–1113. doi: 10.1037/0021-9010.91.5.1102

Bland, J. M., & Altman, D. G. (2000). The odds ratio. *BMJ, 320*, 1468. doi: 10.1136/bmj.320.7247.1468

Blaney, R. L., & Winograd, E. (1978). Developmental differences in children's recognition memory for faces. *Developmental Psychology, 14,* 441–442.

Bornstein, B. H. (1995). Memory processes in elderly eyewitnesses: What we know and what we don't know. *Behavioural Sciences and the Law, 13,* 337–348.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). Comprehensive Meta Analysis: Version 2 [Computer software]. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1,* 97–111. doi: 10.1002/jrsm.12

Bradfield, A. L., Wells, G. L., & Olson, E. A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *Journal of Applied Psychology, 87,* 112–120. doi:10.1037/0021-9010.87.1.112

Brainerd, C. J., & Reyna, V. F. (2004). Fuzzy-trace theory and memory development. *Developmental Review, 24,* 396-439. doi: 10.1016/j.dr.2004.08.005

Brank, E. M. (2007). Elder research: Filling an important gap in psychology and law. *Behavioural Sciences and the Law, 25,* 701–716. doi: 10.1002/bsl.780

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

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

Brooks, N. (1983). *Police Guidelines: Pretrial eyewitness identification procedures.* Law Reform Commission of Canada. Retrieved from http:archive.org/details/policeguidelines00broo

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

*Bruer, K. C., Fitzgerald, R. J., Price, H. L., & Sauer, J. D. (2015). *How sure are you that this is the man you saw? Using confidence judgments to identify a target with child eyewitnesses*. Unpublished manuscript. University of Regina, Canada.

*Cain, W. J., Baker-Ward, L., & Eaton, K. L. (2005). A face in the crowd: The influences of familiarity and delay on preschoolers' recognition. *Psychology, Crime, & Law, 11,* 315–327. doi: 10.1080/10683160418331294835

Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition – A maturational component? *Developmental Psychology, 16,* 257-269.

Ceci, S. J., & Bruck, M. (1993). The suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403-439.

Chance, J. E., Turner, A. L., & Goldstein, A. G. (1982). Development of differential recognition for own- and other-race faces. *Journal of Psychology, 112*, 29–37.

Chung, M., & Thomson, D. M. (1995). Development of face recognition. *British Journal of Psychology, 86,* 55–87.

Clark, S. E. (2012). Costs and benefits of eyewitness identification reform: Psychological science and public policy. *Perspectives on Psychological Science, 7,* 238–259. doi: 10.1177/1745691612439584

*Clifford, B. R. (1993). Witnessing: A comparison of adults and children. In N. K. Clark & G. M. Stephenson (Eds.), *Issues in criminological and legal psychology: Children, evidence and procedure* (pp. 15–21). Leicester, UK: British Psychological Society.

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

*Clifford, B. R., & Toplis, R. (1996). A comparison of adults' and children's witnessing abilities. In N. K. Clark & G. M. Stephenson (Eds.), *Issues in criminological and legal psychology: Investigative and forensic decision making* (pp. 76–83). Leicester, UK: British Psychological Society.

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature-binding difficult? *Developmental Psychology, 42,* 1089-1102. doi: 10.1037/0012-1649.42.6.1089

Craik, F.I., & Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences 10*, 131–138. doi: 1 0.1016/j.tics.2006.01.007

Crookes, K., & McKone, E. (2009). Early maturity of face recognition: No childhood development of holistic processing, novel face encoding, or face-space. *Cognition, 111,* 219–247. doi: 10.1016/j.cognition.2009.02.004

Davies, G. M. (1996). Children's identification evidence. In S. L. Sporer, M. S. Malpass, & G. Koehnken (Eds.) *Psychological issues in eyewitness identification* (pp. 233-258). Mahwah, NJ: Lawrence Erlbaum Associates.

*Davies, G. M., 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). New York: Wiley.

*Davies, G., Tarrant, A., & Flin, R. (1989). Close encounters of the witness kind: Children's memory for a simulated health inspection. *British Journal of Psychology, 80,* 415–429. doi: 10.1111/j.2044-8295.1989.tb02333.x

*Dehon, H., Vanootighem, V., & Brédart, S. (2013). Verbal overshadowing of face memory does occur in children too! *Frontiers in Psychology, 4,* 1-12. doi: 10.3389/fpsyg.2013.00970

*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,* 1–12.

Dent, H. R., & Stephenson, G. M. (1979). Identification evidence: Experimental investigations of factors affecting the reliability of juvenile and adult witnesses. In D. P. Farrington, K.

Hawkins, & S. M. Lloyd-Bostock (Eds.), *Psychology, law and legal processes* (pp. 195–206). Atlantic Highlands, NJ: Humanities Press.

*dePerczel, M. Y. (1991). *The effects of incident involvement on children's eyewitness performance.* Unpublished thesis. West Virginia University, Morgantown, United States.

Devlin, Lord P. (1976). *Report to the Secretary of State for the Home Department of the Departmental Committee on Evidence of Identification in Criminal Cases.* London: Her Majesty's Stationery Office.

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

*Dore, H. S. (2002). *The own-race bias in children's eyewitness memory.* Unpublished thesis. Florida State University, Tallahassee, United States.

Dunlevy, J., & Cherryman, J. (2013). Target-absent eyewitness identification line-ups: Why do children like to choose. *Psychiatry, Psychology and Law, 20*, 1-10. doi: 10.1080/13218719.2012.671584

*Eisen, M. L., Qin, J., Goodman, G. S., & Davis, S. L. (2002). Memory and suggestibility in maltreated children: Age, stress arousal, dissociation, and psychopathology. *Journal of Experimental Child Psychology, 83*, 167–212.

Ellis, H. D., & Flin, R. H. (1990). Encoding and storage effects in 7-year-olds' and 10-year-olds' memory for faces. *British Journal of Developmental Psychology, 8*, 77–92.

Fandakova, Y., Sander, M. C., Werkle-Bergner, M., & Shing, Y. L. (2014). Age differences in short-term memory binding are related to working memory performance across the lifespan. *Psychology and Aging, 29*, 140-149. doi: 10.1037/a0035347

*Fitzgerald, R. J., Gomez, D. M., Connolly, D., & Price, H. (n.d.). *Exploring lineup member similarity effects with children.* Unpublished manuscript. Simon Fraser University, Vancouver, Canada.

*Fitzgerald, R. J., & Price, H. L. (n.d.). *Wildcard effects on children's eyewitness identification.* Unpublished manuscript. University of Regina, Canada.

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

*Fitzgerald, R. J., Whiting, B. F., Therrien, N. M., & Price, H. L. (2014). Lineup member similarity effects on children's eyewitness identification. *Applied Cognitive Psychology*, *28,* 409–418. doi: 10.1002/acp.3012

*Freire, A., Lee, K., Williamson, K. S., Stuart, S. J. E., & Lindsay, R. C. L. (2004). Lineup identification by children: Effects of clothing bias. *Law and Human Behavior, 28,* 339–354. doi: 10.1023/B:LAHU.0000029142.00834.e3

Gallagher, C., Maguire, E. R., Mastrofski, S. D., & Reisig, M. D. (2001). *The public image of the police*. Final report to The International Association of Chiefs of Police by The Administration of Justice Program, George Mason University. Retrieved from http://www.theiacp.org/PoliceServices/ProfessionalAssistance/ThePublicImageofthePolice/tabid/198/Default.aspx#ch2

Gentle, K. (2012). *'I'm pretty sure that might be him'. Investigating the effect of a 'wildcard' on young and older adults' lineup performance.* Unpublished thesis. University of Otago. New Zealand.

*Goetze, H. J. (1980). *The effect of age and method of interview on the accuracy and completeness of eyewitness accounts.* Unpublished thesis. Hofstra University, Long Island, United States.

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

*Goodman, G. S., Hirschman, J., Hepps, D., & Rudy, L. (1991). Children's memory for stressful events. *Merrill-Palmer Quarterly, 37,* 109–158.

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

*Goodsell, C. A. (2006).[7] *Mugshot exposure prior to lineup identification:Commitment, familiarity, and age effects*. Thesis. University of Alabama, Huntsville, Alabama.

Goodsell, C. A., Neuschatz, J. S., & Gronlund, S. D. (2009). Effects of mugshot commitment on lineup performance in young and older adults. *Applied Cognitive Psychology, 23,* 788-803. doi: 10.1002/acp.1512

Gourevitch, V., & Galanter, E. (1967). A significance test for one parameter isosensitivity functions. *Psychometrika, 32,* 25–33. doi: 10.1007/BF02289402

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

[7] Goodsell's (2006) thesis was published; however, the required descriptive statistics were not reported in the published article (Goodsell, Neuschatz, & Gronlund, 2009). Accordingly, the data were extracted from Goodsell's (2006) thesis and the study was coded as published for all analyses involving the publication status covariate.

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

*Hammes, K. M. (1990). *Individual differences in children's eyewitness memory skills.* Unpublished thesis. University of North Dakota, Grand Forks, United States.

*Havard, C. (n.d.). *Can the Mystery Man reduce false identifications for older adult witnesses?* Unpublished manuscript. The Open University, Milton Keynes, England.

*Havard, C., & Memon, A. (2009). The influence of face age on identification from a video lineup: A comparison between older and younger adults. *Memory, 17,* 847–859. doi: 10.1080/09658210903277318

*Havard, C., & Memon, A. (2013). The mystery man can help reduce false identification for child witnesses: Evidence from video line-ups. *Applied Cognitive Psychology, 27,* 50–59. doi: 10.1002/acp.2870

*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. doi: 10.1002/acp.1645

*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,* 929–944. doi: 10.1080/1068316X.2011.598156

Healy, M. R., Light, L. L., & Chung, C. (2005). Dual-process models of associative recognition in young and older adults: Evidence from receiver operating characteristics. *Journal of Experiment Psychology: Learning, Memory and Cognition, 31*, 768-788. doi: 10.1037/0278-7393.31.4.768

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Education Statistics, 6*, 107–128. doi: 10.2307/1164588

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9*, 426–445. doi: 10.1037/1082-989X.9.4.426

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1,* 39–65. doi: 10.1002/jrsm.5

Higgins, J. P. T. (2008). Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. *International Journal of Epidemiology*, *37*, 1158–1160. doi:1158-1160.10.1093/ije/dyn204

Higgins, J. P. T, & Green, S. (2011). Cochrane handbook for systematic reviews of interventions (Version 5.1.0). *The Cochrane Collaboration*. Retrieved from www.cochrane-handbook.org.

*Houston, K. A. (2010). *The emotional eyewitness: An investigation into the effects of anger on eyewitness recall and recognition performance.* Unpublished thesis. University of Aberdeen, Scotland.

*Humphries, J. E., Holliday, R., & Flowe, H. (2012). Faces in motion: Age-related changes in eyewitness identification performance in simultaneous, sequential, and elimination lineups. *Applied Cognitive Psychology, 26,* 149–158. doi: 10.1002/acp.1808

*Huneycutt, D. (2004). *Young eyewitnesses: An examination of young children's response accuracy to target present and target absent lineup arrays following training procedures.* Unpublished thesis. Drexel University, Philadelphia, United States.

Innocence Project. (2014). Understand the causes: Eyewitness misidentification. Retrieved January 12, 2014, from http://www.innocenceproject.org/understand/Eyewitness-Misidentification.php

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513–541.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3-28. doi: 10.1037/0033-2909.114.1.3

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

*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, 102,* 161–183. doi: 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,* 286–314. doi: 10.1016/j.jecp.2007.01.007

*Key, K. N., Cash, D. K., Neuschatz, J. S., Price, J. L., Wetmore, S. A., & Gronlund, S. D. (n.d.). *Showups versus Lineups with Older Adults*. Unpublished manuscript. University of Alabama in Huntsville, University of Oklahoma, and Canisius College (New York). United States.

*King, M. A. (1984). *An investigation of the eyewitness abilities of children.* Unpublished thesis. University of British Columbia, Vancouver, Canada.

*Kinlen, T. J., Adams-Price, C. E., & Henley, T. B. (2007). Verbal overshadowing and face recognition in young and old adults. *Educational Gerontology, 33,* 969–979. doi: 10.1080/03601270701632040

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

Lindsay, 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,* 215-218. doi: 10.1111/1467-9280.00041

*Lindsay, R. C. L, Craig, W., Lee, K., Pozzulo, J. D., Rombough, V., & Smyth, L. (1995, July). *Eyewitness identification procedures for use with children.* Paper presented at the meeting of the Society for Applied Research in Memory and Cognition, Vancouver, British Columbia.

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

Lindsay, R. C. L., & Wells, G. L. (1985). Improving eyewitness identification from lineups: Simultaneous versus sequential lineup presentations. *Journal of Applied Psychology, 70,* 556–564. doi: 10.1037/0021-9010.70.3.556

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. New York, NY: Sage Publications.

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

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and "nonparametric" measures. *Psychological Bulletin, 107,* 401-413. doi: 10.1037/0033-2909.107.3.401

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge: Cambridge University Press.

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

Malpass, R. S., Tredoux, C. G., Schreiber Compo, N., McQuiston-Surrett, D., Maclin, O. H., Zimmerman, L. A., & Topp, L. D. (2008). Study space analysis for policy development. *Applied Cognitive Psychology, 22,* 789-801. doi: 10.1002/acp.1483

Manson v. Braithwaite. (1977). 432 U.S. 98.

*Marin, B. V., Holmes, D. L., Guth, M., & Kovac, P. (1979). The potential of children as eyewitnesses: A comparison of children and adults on eyewitness tasks. *Law and Human Behavior, 3,* 295–306. doi: 10.1007/BF01039808

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology, 38,* 115–142. doi: 10.1037/0012-1649.38.1.115

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory & Cognition, 33,* 783–792. doi: 10.3758/BF03193074

*Memon, A., & Bartlett, J. (2002). The effects of verbalisation on face recognition in young and older adults. *Applied Cognitive Psychology, 16,* 635–650. doi: 10.1002/acp.820

*Memon, A., Bartlett, J., Rose, R. A., & Gray, C. (2003). The aging eyewitness: Effects of age of face, delay, and source memory ability. *The Journal of Gerontology: Psychological Sciences, 58,* 338–345.

*Memon, A., & Gabbert, F. (2003a). Improving the identification accuracy of senior witnesses: Do pre-lineup questions and sequential testing help? *Journal of Applied Psychology, 88,* 341–347. doi: 10.1037/0021-9010.88.2.341

*Memon, A., & Gabbert, F. (2003b). Unravelling the effects of sequential presentation in culprit-present lineups. *Applied Cognitive Psychology, 17*, 703–714. doi: 10.1002/acp.909

*Memon, A., Gabbert, F., & Hope, L. (2004). The ageing eyewitness. In J. Adler (Ed.), *Forensic psychology: Debates, concepts and practice* (pp. 96–112). Uffcolme, Devon: Willan, Forensic Psychology Series.

*Memon, A., Hope, L., Bartlett, J., & Bull, R. (2002). Eyewitness recognition errors: The effects of mugshot viewing and choosing in young and old adults. *Memory and Cognition, 30,* 1219–1227. doi: 10.3758/BF03213404

*Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: Effects on eyewitness accuracy and confidence. *British Journal of Psychology, 94,* 339–354. doi: 10.1348/000712603767876262

*Mertin, P. (1989). The memory of young children for eyewitness events. *Australian Journal of Social Issues, 24,* 23–32.

Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous and sequential lineups. *Journal of Experimental Psychology: Applied, 18,* 361–376. doi: 10.1037/a0030609

Mickes, L., Moreland, M. B., Clark, S. E., & Wixted, J. T. (2014). Missing the information need to perform ROC analysis? Then compute *d'*, not the diagnosticity ratio. *Journal of Applied Research in Memory and Cognition, 3,* 58-62. doi: 10.1016/j.jarmac.2014.04.007

*Miller, C. M. (1993). *The effects of two methods of recall enhancement on child and adult eyewitness memory*. Unpublished thesis. West Virginia University, Morgantown, United States.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31,* 553–566.

*Morten, J. T. P. (2014). *Individual differences in visual and verbal eyewitness testimony: Putting the "I" in eyewitness evidence.* Unpublished thesis. University of Otago, New Zealand.

National Research Council. (2014). *Identifying the culprit: Assessing eyewitness identification*. Washington, DC: The National Academies Press.

Neil v. Biggers. (1972). 409 U.S. 188.

Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant and Child Development, 10,* 3-18. doi: 10.1002/icd.239

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126,* 220–246. doi: 10.1037/0033-2909.126.2.220

*Oates, K., & Shrimpton, S. (1991). Children's memories for stressful and non-stressful events. *Medicine, Science and the Law, 31,* 4–10.

Osborne, J. W. (2006). Bringing balance and technical accuracy to reporting odds ratios and the results of logistic regression analyses. *Practical Assessment Research & Evaluation, 11,* 1-6.

Palmer, M. A., & Brewer, N. (2012). Sequential lineup presentation promotes less biased criterion setting but does not improve discriminability. *Law and Human Behavior, 36,* 247–255. doi: 10.1037/h0093923

Palmer, M. A., Brewer, N., & Weber, N. (2010). Postidentification feedback affects subsequent eyewitness identification performance. *Journal of Experimental Psychology: Applied, 16,* 387–398. doi: 10.1037/a0021034

*Parker, J. F., & Carranza, L. E. (1989). Eyewitness testimony of children in target-present and target-absent lineups. *Law and Human Behavior, 13,* 133–149. doi: 10.1007/BF01055920

*Parker, J. F., Haverfield, E., & Baker-Thomas, S. (1986). Eyewitness testimony of children. *Journal of Applied Social Psychology, 16,* 287–302. doi: 10.1111/j.1559-1816.1986.tb01141.x

*Parker, J. F., & Ryan, V. (1993). An attempt to reduce guessing behavior in children's and adults' eyewitness identifications. *Law and Human Behavior, 17,* 11–26.

doi: 10.1007/BF01044534

Peters, D. P. (1987). The impact of naturally occurring stress on children's memory. In S. J. Ceci, D. F. Ross, & M. P. Toglia (Eds.), *Children's eyewitness memory* (pp. 122-141). New York: Springer-Verlag.

*Pozzulo, J. D., & Balfour, J. (2006). The impact of change in appearance on children's eyewitness identification accuracy: Comparing simultaneous and elimination lineup procedures. *Legal and Criminological Psychology, 11,* 25–34. doi: 10.1348/135532505X52626

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

*Pozzulo, J. D., Dempsey, J. L., 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,* 55–62. doi: 10.1007/s11896-012-9107-5

*Pozzulo, J. D., Dempsey, J. L., Crescini, C., & Lemieux, J. (2009). Examining the relation between recall and recognition in an eyewitness context for children and adults. *Psychology, Crime, & Law, 15,* 409–423. doi: 10.1080/10683160802279625

*Pozzulo, J. D., Dempsey, J. L., & Pettalia, J. (2013). The Z Generation: Examining perpetrator descriptions and lineup identification procedures. *Journal of Police and Criminal Psychology*, *28*, 63–74. doi: 10.1007/s11896-012-9107-5

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

*Pozzulo, J. D., & Lindsay, R. C. L. (1997). Increasing correct identifications by children. *Expert Evidence, 5,* 126–132.

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

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

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

*Price, H. L., & Fitzgerald, R. J. (n.d.a). *Face-Off: A new procedure for child eyewitness identifications*. Unpublished manuscript. University of Regina, Canada.

*Price, H. L., & Fitzgerald, R. J. (n.d.b). *Lineup procedures and lineup member similarity with child witnesses*. Unpublished manuscript. University of Regina, Canada.

R v Turnbull and others. (1976). 63 Cr App R 132.

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin, 140*, 1281-1302. doi: 10.1037/a0037004

*Rochon, C. A. (2014). *Eliminating error with older eyewitnesses using the elimination lineup*. Unpublished thesis. Memorial University, Corner Brook, Canada.

*Rose, R. A., Bull, R., & Vrij, A. (2003). Enhancing older witnesses' identification performance: Context reinstatement is not the answer. *The Canadian Journal of Police and Security Services,* 1, 173–184.

*Rose, R. A., Bull, R., & Vrij, A. (2005). Non-biased lineup instructions do matter: A problem for older witnesses. *Psychology, Crime & Law, 11,* 147–159. doi: 10.1080/10683160512331316307

*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. doi: 10.1007/s10979-006-9034-z

*Rush, E. B., Quas, J. A., Yim, I. S., Clark, S. E., & Larson, R. (2014). Stress, inteviewer support, and children's eyewitness identification accuracy. *Child Development, 85,* 1292–1305. doi: 10.1111/cdev.12177

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science, 13,* 140-144. doi: 10.1111/j.0963-7214.2004.00293.x

Sander, M. C., Lindenberger, U. & Werkle-Bergner, M. (2012). Lifespan age differences in working memory: A two-component framework. *Neuroscience and Biobehavioral Reviews, 36*, 2007-2033. doi: 10.1016/j.neubiorev.2012.06.004

Sander, M. C., Werkle-Bergner, M., & Lindenberger, U. (2011). Building and strategic selection in working memory: A lifespan dissociation. *Psychology and Aging, 26*, 612-624. Doi: 10.1037/a0023055

*Saunders, C. S. (2007). *Cognitive and social aspects of the identification accuracy of child and adult eyewitnesses*. Unpublished thesis. University of Louisiana, Lafayette, United States.

Scammacca, N., Roberts, G., & Stuebing, K. K. (2013). Meta-analysis with complex research designs: Dealing with dependence from multiple measures and multiple group comparisons. *Review of Educational Research, 84,* 328-364. doi:10.3102/0034654313500826.

Schwartz, B. L. (2014). Memory for people: Integration of face, voice, name, and biographical information (pp. 3-19). In T. J. Perfect and D. S. Lindsay (Eds.). *The SAGE handbook of applied memory*. London: SAGE.

*Scogin, F., Calhoon, S., & *D'*Errico, M. (1994). Eyewitness confidence and accuracy among three age cohorts. *Journal of Applied Gerontology*, 13, 172–184.

*Searcy, J. H., Bartlett, J. C., & Memon, A. (1999). Age differences in accuracy and choosing in eyewitness identification and face recognition. *Memory and Cognition*, *27,* 538–552. doi: 10.3758/BF03211547

*Searcy, J., Bartlett, J. C., & Memon, A. (2000). Influence of post event narratives, line-up conditions and individual differences on false identification by young and older eyewitnesses. *Legal and Criminological Psychology, 5,* 219–235. doi: 10.1348/135532500168100

Searcy, J. H., Bartlett, J. C., Memon, A., & Swanson, K. (2001). Aging and lineup performance at long retention intervals: Effect of metamemory and context reinstatement. *Journal of Applied Psychology, 86,* 207-214. doi: 10.1037//0021-9010.86.2.207

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

Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Muller, V., Li, S.-C., & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience and Biobehavioral Reviews, 34,* 1080-1091. doi: 10.1016/j.neubiorev.2009.11.002

Shing, Y. L., Werkle-Bergner., M., Li, S.-C., & Lindenberger, U. (2008). Associative and strategic components of episodic memory: A lifespan dissociation. *Journal of Experimental Psychology: General, 137*, 495-513. doi: 10.1037/0096-3445.137.3.495

Sinopoli, K. J., Schachar, R., & Dennis, M. (2011). Reward improves cancellation and restraint inhibition across childhood and adolescence. *Developmental Psychology, 47,* 1479–1489. doi: 10.1037/a0024440

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34-50. doi: 10.1037/0096-3445.117.1.34

*Soppe, H. J. G. (1986). Children's recognition of unfamiliar faces: Developments and determinants. *International Journal of Behavioral Development, 9,* 219–233. doi: 10.1177/016502548600900206

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging, 10*, 527–539.

Sporer, S. L., & Marchuk, N. (2014). The reliability of eyewitness identifications by the elderly: An evidence-based review. In M. P. Toglia, D. F. Ross, J. Pozzulo, & E. Pica (Eds.), *The elderly eyewitness in court* (pp. 3-37). New York: Psychology Press.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*, 137–149.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2003). Eyewitness accuracy rates in police showup and lineup presentations: A meta-analytic comparison. *Law and Human Behavior, 27,* 523–540. doi: 10.1023/A:1025438223608

Tanaka, J. W., & Farah, M. J. (1993) Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46,* 225–245. doi: 10.1080/14640749308401045

Tanner-Smith, E. E., & Tipton, E. (2013). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods, 5,* 13-30. doi: 10.1002/jrsm.1091.

*Therrien, N. M. (2012). *Practice does not make perfect in children's eyewitness identification accuracy*. Unpublished thesis. University of Regina, Canada.

Tipton, E. (2014). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods*. Advance online publication. doi: 10.1037/met0000011

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

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: a meta-analysis of training studies. *Psychological Bulletin, 139*, 352–402. doi: 10.1037/a0028446

*von Zedlitz-Neukirch, S. F. (2013). *Children as witnesses: The age effect on the correctness and metacognition in sequential and elimination line-up* [translated from Swedish]. Unpublished thesis. Lund University, Sweden.

Wells, G. L. (1978). Applied eyewitness testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36*, 1546–1557. doi: 10.1037/0022-3514.36.12.1546

Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology, 14,* 89–103. doi: 10.1111/j.1559-1816.1984.tb02223.x

Wells, G. L., & Penrod, S. D. (2011). Eyewitness identification research: Strengths and weaknesses of alternative methods. In B. Rosenfeld, & S. D. Penrod (Eds.), *Research methods in forensic psychology (pp. 237–256).*. Hoboken, NJ: John Wiley and Sons.

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

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

*Wilcock, R. A., & Bull, R. (2010). Novel lineup methods for improving the performance of older eyewitnesses. *Applied Cognitive Psychology, 24,* 718–736. doi: 10.1002/acp.1582

*Wilcock, R. A., Bull, R., & Vrij, A. (2005). Aiding the performance of older eyewitnesses: Enhanced non-biased lineup instructions and lineup presentation. *Psychiatry, Psychology, and Law, 12,* 129–140. doi: 10.1375/pplt.2005.12.1.129

*Wilcock, R. A., Bull, R., & Vrij, A. (2007). Are older witnesses always poorer witnesses? Identification accuracy, context reinstatement, own age bias. *Psychology, Crime and Law 13,* 305–316. doi: 10.1080/10683160600822212

Wixted, J. T., & Mickes, L. (2012). The field of eyewitness memory should abandon probative value and embrace receiver operator characteristic analysis. *Perspective on Psychology Science, 7,* 275-278. doi: 10.1177/1745691612442906

Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychological Review, 121*, 262–276. doi: 10.1037/a0035940

*Wright, D. B., & Stroud, J. S. (2002). Age differences in line-up identification accuracy: People are better with their own age. *Law and Human Behavior, 26,* 641–654. doi: 10.1023/A:1020981501383

Wylie, L. E., Bergt, S., Haby, J., Brank, E. M., & Bornstein, B. H. (2014). Age and lineup type differences in the own-race bias. *Psychology, Crime & Law*. Advance online publication. doi: 10.1080/1068316X.2014.989173

Yarmey, A. D. (1988). Streetproofing and bystander's memory for a child abduction. In M. Gruneberg, P. Morris, & R. Sykes (Eds.), *Practical aspects of memory: Current research and issues, Vol. 1: Memory in everyday life* (pp. 112-116). Chichester, England: Wiley.

*Yarmey, A. D., Jones, H. T., & Rashid, S. (1984). Eyewitness memory of elderly and young adults. In D. J. Muller, D. E. Blackman, & A. J. Chapman (Eds.), *Psychology and law* (pp. 215–228). Chichester, UK: John Wiley and Sons.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441-517. doi: 10.1006/jmla.2002.2864

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

Table 1

Young adults versus children/older adults: Age differences in response rates

Age Groups Compared						Weighted Means		Effect Size & CI[95]			Test of Null			Heterogeneity	
Outcome	Group 1	Group 2	Outliers	*m*	*k*	Group 1	Group 2	ES	*LL*	*UL*	*t*	*df*	*p*	τ^2	I^2
Hits	Young Adult	Child (4-17 yrs)	9	30	107	.55	.47	1.42	1.20	1.69	4.33	21	.001	0.07	17.7
		3-4 yrs	-	4	5	.58	.40	-	-	-	-	-	-	-	-
		5-8 yrs	4	15	33	.57	.49	1.51	1.04	2.18	2.39	12	.034	0.14	23.2
		9-13 yrs	7	20	61	.56	.51	1.22	1.01	1.47	2.28	11	.044	0.02	7.5
		14-17 yrs	-	3	6	.46	.48	-	-	-	-	-	-	-	-
		Older Adult (45-77 yrs)	3	23	59	.48	.38	1.71	1.37	2.14	5.03	18	.001	0.15	30.1
		45-48 yrs	-	3	8	.48	.41	-	-	-	-	-	-	-	-
		68-77 yrs	4	21	50	.52	.38	1.95	1.54	2.48	5.97	15	.001	0.04	10.7
Filler Selections	Young Adult	Child (4-17 yrs)	10	23	87	.22	.31	1.72	1.34	2.20	4.66	13	.001	0.06	12.6
		3-4 yrs	-	2	2	.26	.33	-	-	-	-	-	-	-	-
		5-8 yrs	5	13	25	.22	.33	1.89	1.12	3.13	2.74	10	.021	0.23	29.5
		9-13 yrs	4	16	55	.23	.32	1.71	1.25	2.33	4.23	6	.006	0.01	2.7
		14-17 yrs	-	3	6	.23	.24	-	-	-	-	-	-	-	-
		Older Adult (45-77 yrs)	2	21	55	.25	.40	2.37	1.68	3.35	5.28	18	.001	0.41	50.4
		45-48 yrs	-	3	8	.39	.43	-	-	-	-	-	-	-	-

		68-77 yrs	1	19	48	.21	.38	2.86	2.04	4.00	6.59	16	.001	0.33	39.0
Incorrect Rejects	Young Adult	Child (4-17 yrs)	5	20	86	.31	.27	1.26	1.03	1.54	2.48	12	.029	0.05	12.0
		3-4 yrs	-	2	2	.17	.24	-	-	-	-	-	-	-	-
		5-8 yrs	3	10	23	.29	.27	1.21	0.73	2.01	0.89	7	.403	0.17	21.4
		9-13 yrs	2	15	55	.29	.26	1.23	0.96	1.58	1.88	9	.092	0.05	11.9
		14-17 yrs	-	3	6	.31	.30	-	-	-	-	-	-	-	-
		Older Adult (45-77 yrs)	10	20	47	.29	.27	1.06	0.89	1.25	0.71	12	.492	0.00	0.0
		45-48 yrs	-	3	8	.18	.20	-	-	-	-	-	-	-	-
		68-77 yrs	2	19	47	.32	.28	1.21	0.86	1.68	1.19	15	.252	0.21	33.8
Correct Rejects	Young Adult	Child (4-17 yrs)	8	26	68	.57	.43	2.20	1.55	2.66	5.50	20	.001	0.16	32.6
		3-4 yrs	-	1	1	.41	.16	-	-	-	-	-	-	-	-
		5-8 yrs	2	8	13	.64	.41	2.75	1.70	4.44	5.27	6	.002	0.15	29.3
		9-13 yrs	4	20	50	.57	.43	2.04	1.44	2.89	4.38	16	.001	0.19	37.1
		14-17 yrs	-	3	6	.48	.44	-	-	-	-	-	-	-	-
		Older Adult (45-77 yrs)	5	20	47	.54	.38	2.03	1.52	2.71	5.22	16	.001	0.11	26.7
		45-48 yrs	-	2	4	.37	.36	-	-	-	-	-	-	-	-
		68-77 yrs	5	19	43	.55	.34	2.14	1.58	2.86	5.53	14	.001	0.09	21.7
Choosing	Young Adult	Child (4-17 yrs)	8	17	55	.56	.63	1.38	1.20	1.61	4.93	10	.001	0.01	8.1
		3-4 yrs	-	1	1	.68	.82	-	-	-	-	-	-	-	-

		5-8 yrs	0	8	14	.54	.66	1.72	0.88	3.33	1.93	7	.095	0.66	79.2
		9-13 yrs	3	13	39	.55	.62	1.39	1.11	1.75	3.28	9	.010	0.04	24.0
		14-17 yrs	-	3	6	.67	.69	-	-	-	-	-	-	-	-
		Older Adult (45-77 yrs)	4	17	37	.60	.70	1.56	1.23	1.98	4.02	14	.001	0.09	35.8
		45-48 yrs	-	2	4	.71	.72	-	-	-	-	-	-	-	-
		68-77 yrs	1	16	36	.58	.69	1.69	1.27	2.27	3.93	14	.001	0.27	62.8
Sensitivity	Young Adults	Child (4-17 yrs)	3	22	68	1.69	1.31	0.11	0.07	0.15	5.73	13	.001	0.00	0.0
		3-4 yrs	-	1	1	1.26	0.97	-	-	-	-	-	-	-	-
		5-8 yrs	2	8	12	1.76	1.34	0.12	0.01	0.24	2.52	6	.044	0.00	0.0
		9-13 yrs	1	17	49	1.74	1.34	0.11	0.07	0.15	5.91	9	.001	0.00	0.0
		14-17 yrs	-	3	6	1.32	1.30	-	-	-	-	-	-	-	-
		Older Adult (45-77 yrs)	1	18	42	1.54	0.95	0.18	0.13	0.23	8.26	12	.001	0.00	0.0
		45-48 yrs	-	2	4	1.47	1.26	-	-	-	-	-	-	-	-
		68-77 yrs	0	17	39	1.63	1.01	0.19	0.14	0.24	8.34	10	.001	0.00	0.0
Diagnosticity	Young Adults	Child (4-17 yrs)	1	22	70	9.36	5.41	1.69	1.37	2.07	5.55	12	.001	0.00	0.0
		3-4 yrs	-	1	1	4.95	3.27	-	-	-	-	-	-	-	-
		5-8 yrs	0	8	14	10.16	5.69	1.95	1.07	3.60	2.82	5	.035	0.00	0.0
		9-13 yrs	1	17	49	9.52	5.55	1.59	1.26	2.01	4.56	8	.002	0.00	0.0
		14-17 yrs	-	3	6	6.00	5.36	-	-	-	-	-	-	-	-

		Older Adult (45-77 yrs)	0	18	43	7.61	4.31	1.77	1.30	2.44	4.02	10	.003	0.00	0.0
		45-48 yrs	-	2	4	7.18	5.51	-	-	-	-	-	-	-	-
		68-77 yrs	0	17	39	8.34	4.54	1.90	1.40	2.59	4.88	8	.001	0.00	0.0
Suspect Bias	Young Adults	Child (4-17 yrs)	2	21	69	0.74	0.75	-0.01	-0.05	0.03	0.40	12	.726	0.00	0.0
		3-4 yrs	-	1	1	0.66	0.60	-	-	-	-	-	-	-	-
		5-8 yrs	2	7	12	0.64	0.75	-0.06	-0.15	0.03	1.70	5	.147	0.00	0.00
		9-13 yrs	0	17	50	0.73	0.74	-0.01	-0.06	0.05	0.21	9	.838	0.00	0.00
		14-17 yrs	-	3	6	0.75	0.71	-	-	-	-	-	-	-	-
		Older Adult (45-77 yrs)	0	18	43	0.74	0.83	-0.05	-0.10	-0.01	2.55	12	.026	0.00	0.0
		45-48 yrs	-	2	4	0.59	0.68	-	-	-	-	-	-	-	-
		68-77 yrs	0	17	39	0.69	0.81	-0.07	-0.14	-0.01	2.34	10	.040	0.00	0.0

Note. CI = confidence interval; ES = effect size; *LL* = lower limit; *UL* = upper limit. For the five lineup response outcomes (hits, filler selections, incorrect rejects, correct rejects, and choosing), the weighted means are proportions and the effect size is odds ratio. For sensitivity, the weighted means are *d'* scores and the effect size is Hedges' *g*. For diagnosticity, the weighted means are relative risks and the effect size is ratio of relative risks. For suspect bias, the weighted means are $c_{suspect}$ scores and the effect size is Hedges' *g*. All statistics depicted are with outliers removed (Outliers = number of *k* removed).

Table 2

Older children versus younger children: Age differences in response rates, sensitivity, diagnosticity, and suspect bias

				Weighted Means		Effect Size & CI[95]			Test of Null			Heterogeneity	
Outcome	Outliers	*m*	*k*	Older Children	Younger Children	ES	*LL*	*UL*	*t*	*df*	*p*	τ^2	I^2
Hits	18	40	153	.53	.46	1.46	1.22	1.75	4.26	27	.001	0.07	13.8
Filler Selections	12	29	114	.25	.34	1.64	1.32	2.00	4.88	19	.001	0.00	0.0
Incorrect Rejects	11	30	120	.24	.29	1.27	1.06	1.54	2.69	22	.013	0.02	2.3
Correct Rejects	8	27	103	.51	.46	1.23	0.97	1.57	1.81	22	.083	0.17	26.8
Choosing	12	23	77	.63	.61	1.06	0.90	1.26	0.81	19	.428	0.00	0.0
Sensitivity	0	26	92	1.65	1.46	0.05	0.01	0.11	2.28	22	.033	0.00	0.0
Diagnosticity	0	26	92	8.42	7.16	1.22	1.00	1.48	2.13	16	.049	0.00	0.0
Suspect Bias	1	26	91	0.67	0.73	-0.03	-0.08	0.01	1.58	22	.128	0.00	0.0

Note. CI = confidence interval; ES = effect size, *LL* = lower limit; *UL* = upper limit. For the five lineup response outcomes (hits, filler selections, incorrect rejects, correct rejects, and choosing), the weighted means are proportions and the effect size is odds ratio. For sensitivity, the weighted means are *d'* scores and the effect size is Hedges' *g*. For diagnosticity, the weighted means are relative risks and the effect size is ratio of relative risks. For suspect bias, the weighted means are $c_{suspect}$ scores and the effect size is Hedges' *g*. All statistics depicted are with outliers removed (Outliers = number of *k* removed).

Table 3

Regression coefficients for mean age and mean age difference covariates

				Coefficient & CI[95]			Test of Null		
Dataset	Covariate	Outcome	Effect Type	Estimate	*LL*	*UL*	*t*	*df*	*p*
Young Adults vs. Children	Mean Age (Child group)	Hits	Within-Study	0.94	0.87	1.02	1.92	5	.108
			Between-Study	0.95	0.90	1.01	2.08	8	.074
			Both	0.95	0.90	0.99	2.63	10	.025
		Filler Selections	Within-Study	1.06	0.96	1.17	1.78	4	.152
			Between-Study	1.03	0.93	1.14	0.76	6	.480
			Both	1.04	0.97	1.12	1.23	8	.253
		Incorrect Rejects	Within-Study	0.97	0.90	1.05	0.90	4	.415
			Between-Study	0.99	0.90	1.09	0.13	4	.901
			Both	0.99	0.93	1.05	0.37	6	.723
		Correct Rejects	Both	0.97	0.90	1.04	0.87	7	.412
		Choosing	Both	1.01	0.97	1.05	0.68	6	.524
		Sensitivity	Both	-0.01	-0.04	0.02	0.77	8	.464
		Diagnosticity	Both	0.97	0.88	1.06	0.77	7	.468
		Suspect Bias	Both	0.01	0.00	0.02	1.97	7	.090
Older vs. Younger Children	Mean Age Difference	Hits	Within	1.12	1.03	1.20	3.34	6	.017

	Between	1.02	0.87	1.21	0.30	15	.770
Filler Selections	Within	0.89	0.78	1.01	2.78	4	.056
	Between	1.02	0.83	1.25	0.18	10	.861
Incorrect Rejects	Within	1.12	0.83	1.51	0.99	4	.376
	Between	1.00	0.89	1.13	0.04	11	.968
Correct Rejects	Within	1.01	0.88	1.15	0.18	5	.867
	Between	1.08	0.84	1.40	0.68	9	.518
Choosing	Both	0.95	0.87	1.04	1.29	7	.237
Sensitivity	Both	0.03	-0.01	0.06	1.79	10	.104
Diagnosticity	Both	1.13	1.01	1.27	2.41	7	.045
Suspect Bias	Both	0.01	-0.03	0.04	0.35	10	.733

Note. CI = confidence interval; *LL* = lower limit; *UL* = upper limit. For the five lineup response outcomes (hits, filler selections, incorrect rejects, correct rejects, and choosing), the estimate is the ratio of odds ratios. For sensitivity and suspect bias, the estimate is the difference between Hedges' *g* scores. For diagnosticity, the estimate is the ratio of ratios of relative risks.

Table 4

Regression coefficients for publication status and publication year covariates

			Coefficient & CI[95]			Test of Null		
Dataset	Outcome	Covariate	Estimate	*LL*	*UL*	*t*	*df*	*p*
Young Adult vs Child/Older Adult	Hits	Publication Year	1.01	0.99	1.03	0.89	16	.385
		Publication Status	1.13	0.78	1.62	0.69	15	.504
	Filler Selections	Publication Year	0.99	0.96	1.02	0.67	11	.515
		Publication Status	0.63	0.39	1.02	2.11	12	.058
	Incorrect Rejects	Publication Year	0.99	0.97	1.01	1.09	4	.336
		Publication Status	1.04	0.72	1.52	0.24	11	.812
	Correct Rejects	Publication Year	0.98	0.95	1.00	2.63	7	.032
		Publication Status	1.38	0.95	1.97	1.90	10	.085
	Choosing	Publication Year	1.02	1.01	1.04	4.12	5	.010
		Publication Status	0.96	0.77	1.19	0.45	9	.662
	Sensitivity	Publication Year	0.00	-0.01	0.01	0.48	6	.650
		Publication Status	0.09	-0.01	0.18	1.95	11	.076
	Diagnosticity	Publication Year	1.00	0.97	1.04	0.24	10	.819
		Publication Status	1.54	1.07	2.20	2.67	11	.022
	Suspect Bias	Publication Year	0.00	-0.01	0.00	1.77	6	.127

		Publication Status	-0.02	-0.09	0.05	0.50	11	.629
Older Children vs. Younger Children	Hits	Publication Year	1.00	0.98	1.02	0.14	14	.889
		Publication Status	1.15	0.77	1.72	0.72	17	.482
	Filler Selections	Publication Year	1.01	0.99	1.03	0.77	7	.468
		Publication Status	0.71	0.46	1.09	1.76	10	.109
	Incorrect Rejects	Publication Year	1.00	0.97	1.02	0.32	7	.761
		Publication Status	0.90	0.62	1.34	0.54	13	.601
	Correct Rejects	Publication Year	0.96	0.93	0.98	3.42	12	.005
		Publication Status	0.99	0.64	1.52	0.08	9	.940
	Choosing	Publication Year	1.02	0.99	1.04	1.91	8	.091
		Publication Status	1.07	0.78	1.48	0.48	11	.640
	Sensitivity	Publication Year	0.00	-0.01	0.01	1.12	11	.286
		Publication Status	-0.05	-0.16	0.06	0.92	15	.372
	Diagnosticity	Publication Year	0.99	0.97	1.01	1.04	11	.320
		Publication Status	0.84	0.56	1.28	0.87	13	.401
	Suspect Bias	Publication Year	0.00	-0.01	0.01	0.05	11	.959
		Publication Status	0.04	-0.06	0.13	0.78	15	.448

Note. CI = confidence interval; *LL* = lower limit; *UL* = upper limit. For the five lineup response outcomes (hits, filler selections, incorrect rejects, correct rejects, and choosing), the estimate is the ratio of odds ratios. For sensitivity and suspect bias, the estimate is the difference between Hedges' *g* scores. For diagnosticity, the estimate is the ratio of ratios of relative risks.

Table 5

Young adults versus children: Procedural influences on age differences (within-study effects)

					Weighted Means		Effect Size & CI95			Moderator Test		
Procedures Compared	Outcome	Procedure	*m*	*k*	Young Adults	Children	ES	*LL*	*UL*	*Q*	*df*	*p*
Lineup vs. Showup	Hits	Lineup	3	5	.53	.64	1.64	1.01	2.68	1.53	1	.216
		Showup	3	3	.49	.77	3.55	1.70	7.45			
	Correct Rejects	Lineup	2	3	.70	.30	5.37	2.89	9.99	1.23	1	.267
		Showup	2	2	.94	.62	9.84	4.33	22.38			
	Sensitivity	Lineup	2	3	1.60	1.83	0.08	-0.10	0.26	0.62	1	.433
		Showup	2	2	1.37	1.03	0.07	-0.17	0.29			
	Diagnosticity	Lineup	2	3	10.03	6.27	1.50	0.53	4.19	2.13	1	.145
		Showup	2	2	17.39	8.19	3.87	1.81	8.27			
	Suspect Bias	Lineup	2	3	0.82	0.25	0.34	0.15	0.53	4.18	1	.041
		Showup	2	2	0.88	-0.18	0.72	0.50	0.95			
Elimination vs. Simultaneous	Hits	Elimination	4	5	.51	.48	1.14	0.71	1.83	0.44	1	.507
		Simultaneous	4	5	.64	.57	1.44	0.89	2.32			
	Filler Selections	Elimination	4	5	.11	.27	3.22	1.55	6.69	0.18	1	.668
		Simultaneous	4	5	.19	.35	2.39	1.32	4.33			

	Incorrect Rejects	Elimination	4	5	.45	.35	1.61	1.01	2.58	0.00	1	.973
		Simultaneous	4	5	.23	.16	1.59	0.87	2.90			
	Correct Rejects	Elimination	4	5	.77	.55	2.92	1.72	4.95	0.21	1	.646
		Simultaneous	4	5	.61	.45	2.17	1.33	3.54			
	Choosing	Elimination	4	5	.38	.54	1.96	1.42	2.71	0.15	1	.698
		Simultaneous	4	5	.57	.69	1.76	1.25	2.46			
	Sensitivity	Elimination	4	5	1.79	1.31	0.13	-0.03	0.28	0.01	1	.910
		Simultaneous	4	5	1.98	1.49	0.14	-0.02	0.29			
	Diagnosticity	Elimination	4	5	15.37	5.77	2.33	0.76	7.12	0.24	1	.622
		Simultaneous	4	5	11.37	5.92	1.65	0.73	3.73			
	Suspect Bias	Elimination	4	5	0.98	0.83	0.08	-0.08	0.23	0.38	1	.539
		Simultaneous	4	5	0.61	0.59	0.01	-0.15	0.16			
Sequential vs. Simultaneous	Hits	Sequential	6	7	.56	.45	1.53	0.96	2.41	0.43	1	.508
		Simultaneous	6	7	.59	.60	1.04	0.65	1.66			
	Filler Selections	Sequential	3	4	.21	.36	2.19	0.92	5.24	0.54	1	.462
		Simultaneous	3	4	.19	.25	1.42	0.70	2.88			
	Incorrect Rejects	Sequential	3	4	.30	.34	0.80	0.40	1.61	0.47	1	.492
		Simultaneous	3	4	.28	.16	2.00	0.92	4.38			
	Correct Rejects	Sequential	5	6	.73	.33	5.79	3.38	9.93	6.36	1	.012

	Simultaneous	5	6	.55	.40	1.89	1.56	3.08			
Choosing	Sequential	3	4	.50	.65	1.91	1.25	2.91	0.70	1	.402
	Simultaneous	3	4	.62	.71	1.52	0.98	2.36			
Sensitivity	Sequential	5	6	1.73	1.03	0.21	0.05	0.36	2.48	1	.115
	Simultaneous	5	6	1.64	1.58	0.02	-0.15	0.18			
Diagnosticity	Sequential	5	6	12.15	3.79	2.74	1.05	7.15	1.38	1	.240
	Simultaneous	5	6	7.70	6.18	1.29	0.58	2.90			
Suspect Bias	Sequential	5	6	0.81	0.68	0.07	-0.09	0.23	0.01	1	.951
	Simultaneous	5	6	0.63	0.49	0.09	-0.08	0.25			

Note. Note. CI = confidence interval; ES = effect size; *LL* = lower limit; *UL* = upper limit. For the five lineup response outcomes (hits, filler selections, incorrect rejects, correct rejects, and choosing), the weighted means are proportions and the effect size is odds ratio. For sensitivity, the weighted means are *d'* scores and the effect size is Hedges' *g*. For diagnosticity, the weighted means are relative risks and the effect size is ratio of relative risks. For suspect bias, the weighted means are $c_{suspect}$ scores and the effect size is Hedges' *g*.

Table 6

Differences between young adults and children (4-17 years) for all studies with simultaneous presentation

			Weighted Means		Effect Size & CI[95]			Test of Null		
Outcome	*m*	*k*	Young Adults	Children	ES	*LL*	*UL*	*t*	*df*	*p*
Hits	29	72	.57	.51	1.44	1.06	1.96	2.42	26	.023
Filler Selections	21	59	.23	.32	1.77	1.24	2.53	3.39	16	.004
Incorrect Rejects	18	51	.26	.22	1.24	0.87	1.79	1.29	14	.217
Correct Rejects	24	51	.55	.40	2.11	1.53	2.92	4.88	19	.001
Choosing	16	41	.59	.68	1.54	1.16	2.07	3.23	14	.006
Sensitivity	20	46	1.73	1.37	0.11	0.05	0.17	3.83	11	.002
Diagnosticity	20	46	8.60	5.39	1.54	1.24	1.89	4.59	9	.001
Suspect Bias	20	46	0.66	0.66	0.00	-0.06	0.06	0.06	11	.953

Note. CI = confidence interval; ES = effect size; *LL* = lower limit; *UL* = upper limit. For the five lineup response outcomes (hits, filler selections, incorrect rejects, correct rejects, and choosing), the weighted means are proportions and the effect size is odds ratio. For sensitivity, the weighted means are *d'* scores and the effect size is Hedges' *g*. For diagnosticity, the weighted means are relative risks and the effect size is ratio of relative risks. For suspect bias, the weighted means are $c_{suspect}$ scores and the effect size is Hedges' *g*.

Table 7

Regression coefficients for young adult comparison group (children vs. older adults)

	Coefficient & CI[95]			Test of Null		
Outcome	Estimate	*LL*	*UL*	*t*	*df*	*p*
Hits	1.20	0.91	1.57	1.35	37	.187
Filler Selections	0.76	0.49	1.17	1.29	33	.205
Incorrect Rejects	0.84	0.68	1.03	1.74	21	.096
Correct Rejects	1.00	0.68	1.48	0.01	34	.990
Choosing	0.90	0.68	1.19	0.83	25	.415
Sensitivity	0.07	0.01	0.13	2.30	23	.031
Diagnosticity	1.04	0.71	1.52	0.22	20	.830
Suspect Bias	-0.05	-0.10	0.01	1.76	22	.093

Note. CI = confidence interval; ES = effect size; *LL* = lower limit; *UL* = upper limit. For the five lineup response outcomes (hits, filler selections, incorrect rejects, correct rejects, and choosing), the estimate is the ratio of odds ratios. For sensitivity and suspect bias, the estimate is the difference between Hedges' *g* scores. For diagnosticity, the estimate is the ratio of ratios of relative risks.

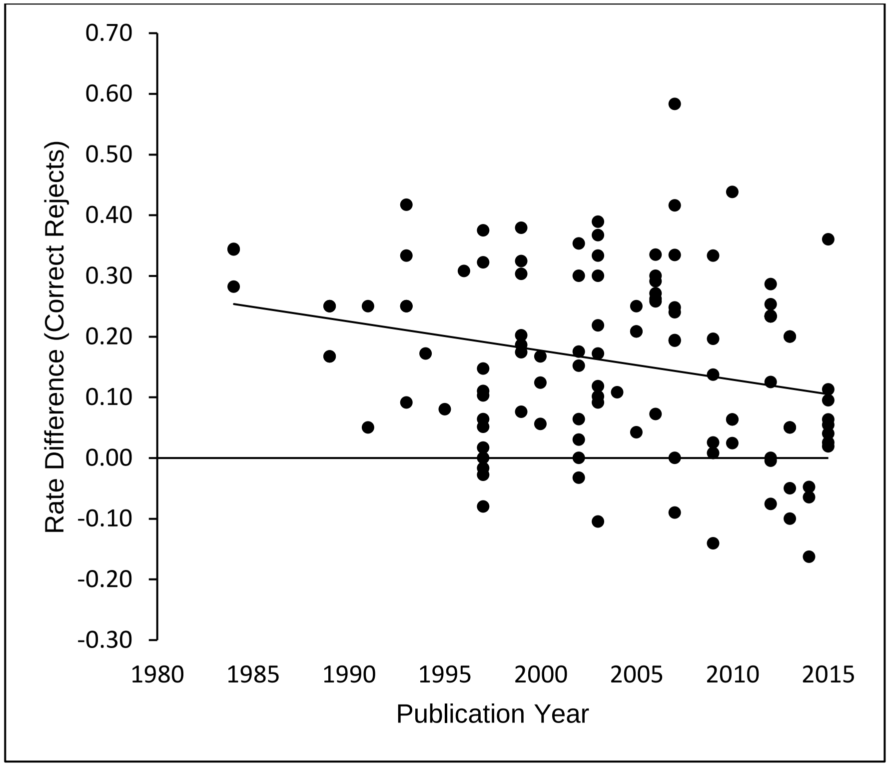

Figure 1. Age differences (young adults vs. children/older adults) in correct reject rates as a function of publication year. Each data point represents a rate difference, which was the child/older adult rate subtracted from the young adult rate.

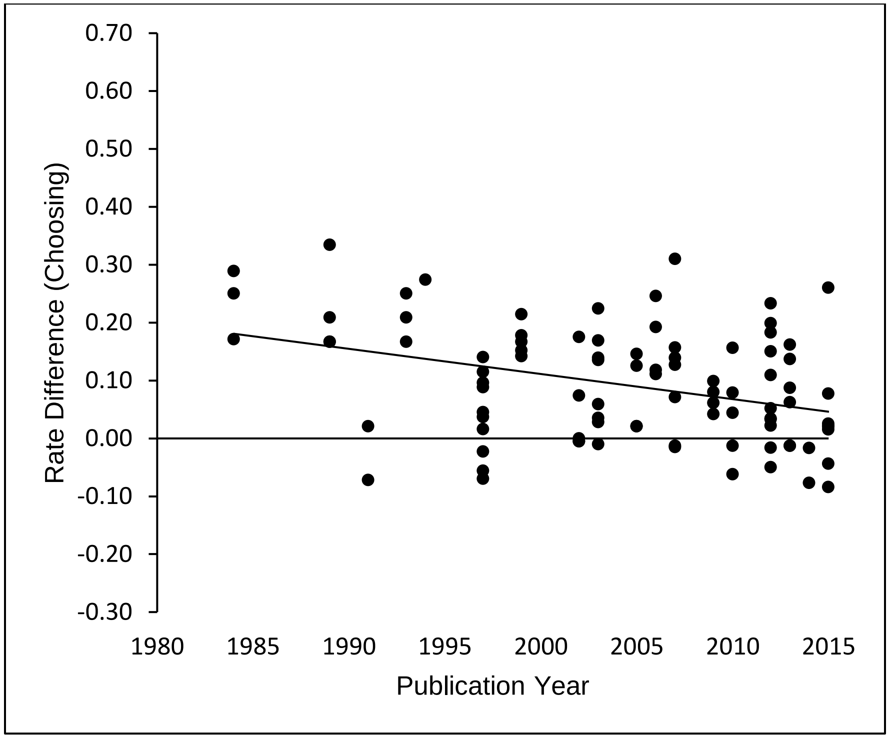

Figure 2. Age differences (young adults vs. children/older adults) in choosing rates as a function of publication year. Each data point represents a rate difference, which was the young adult rate subtracted from the child/older adult rate.

Appendix A

Table A1

Lineup response rates for comparisons between young adults and children

			Young Adults							Children						
				Culprit Present			Culprit Absent				Culprit Present			Culprit Absent		
Year	Authors (Experiment)	Condition	Age	Hit	Filler	Reject	FA	Filler	Reject	Age	Hit	Filler	Reject	FA	Filler	Reject
1979	Marin et al.	Male witness	NR	.50	.50	N/A	-	-	-	6	.41	.59	N/A	-	-	-
										9	.41	.59	N/A	-	-	-
										13	.67	.33	N/A	-	-	-
		Female witness	NR	.58	.42	N/A	-	-	-	6	.67	.33	N/A	-	-	-
										9	.50	.50	N/A	-	-	-
										13	.83	.17	N/A	-	-	-
1986	Goodman & Reed	N/A	22	.75	.19	.06	-	-	-	4	.38	.31	.31	-	-	-
										7	.95	.05	.00	-	-	-
1986	Parker et al.	N/A	24	.58	.42	N/A	-	-	-	8	.71	.29	N/A	-	-	-
1989	Mertin	N/A	NR	.95	.05	N/A	-	-	-	5	.60	.40	N/A	-	-	-
										13	.68	.33	N/A	-	-	-
1989	Parker et al.	Adult target 1	21	.08	.25	.67	.06	.28	.67	9	.33	.42	.25	.10	.49	.42
		Adult target 2	21	.25	.42	.33	.06	.28	.67	9	.42	.33	.25	.10	.49	.42
		Child target 1	21	.42	.25	.33	.13	.63	.25	9	.50	.42	.08	.15	.76	.08
		Child target 2	21	.50	.25	.25	.13	.63	.25	9	.33	.58	.08	.15	.76	.08
1991	Lieppe	Toucher	20	.93	.00	.07	.01	.07	.92	6	.79	.00	.21	.06	.28	.67
										10	.63	.19	.19	.02	.11	.87
		Intruder	20	.81	.00	.19	-	-	-	6	.38	.23	.38	-	-	-

										10	.45	.23	.32	-	-	-
1991	Goodman, Hirschman et al.	Parents vs. younger kids	NR	.58	NR	NR	-	-	-	4	.30	NR	NR	-	-	-
		Parents vs. older kids	NR	.44	NR	NR	-	-	-	6	.54	NR	NR	-	-	-
1993	Clifford	N/A	28	.44	NR	NR	.00	.00	1.00	5	.18	NR	NR	NR	NR	.91
1993	Parker et al.	Simultaneous, control	24	.42	.17	.42	.10	.49	.42	9	.42	.33	.25	.14	.69	.17
		Simultaneous, practice	24	.33	.17	.50	.04	.21	.75	9	.42	.33	.25	.08	.42	.50
		Sequential, control	24	.08	.17	.75	.04	.21	.75	9	.25	.58	.17	.11	.56	.33
		Sequential, practice	24	.50	.33	.17	.07	.35	.58	9	.33	.50	.17	.13	.63	.25
1993	Miller	Cognitive Interview	NR	.88	NR	NR	-	-	-	8	.69	NR	NR	-	-	-
		Visual props	NR	.75	NR	NR	-	-	-	8	.38	NR	NR	-	-	-
		baseline	NR	.88	NR	NR	-	-	-	8	.69	NR	NR	-	-	-
1995	Lindsay et al.	Simultaneous	NR	.58	NR	NR	.11	.56	.33	10	.25	NR	NR	.13	.63	.25
		Sequential	NR	.00	NR	NR	.01	.07	.92	10	.17	NR	NR	.15	.77	.08
1996	Clifford & Toplis	Female Target	NR	.50	.22	.28	-	-	-	6	.74	.26	.00	-	-	-
										9	.22	.17	.61	-	-	-
										12	.39	.39	.22	-	-	-
		Male Target	NR	.17	.50	.33	-	-	-	6	.16	.26	.58	-	-	-
										9	.00	.44	.56	-	-	-
										12	.11	.28	.61	-	-	-
1996	Dekle et al.	N/A	NR	.30	.13	.57	.05	.25	.70	6	.61	.33	.06	.10	.51	.39
1997	Lindsay et al. (1)	Sequential	NR	.62	NR	NR	.04	.21	.75	10	.65	NR	NR	.13	.65	.21
										12	.71	NR	NR	.13	.67	.20
		Simultaneous	NR	.55	NR	NR	.06	.29	.66	10	.71	NR	NR	.12	.60	.28
										12	.80	NR	NR	.11	.56	.33
1997	Lindsay et al. (2)	Sequential	NR	.45	NR	NR	-	-	-	4	.26	NR	NR	-	-	-

		Simultaneous	NR	.80	NR	NR	-	-	-	4	.53	NR	NR	-	-	-
1997	Pozzulo & Lindsay	Control	19	.52	.08	.40	.10	.48	.43	10	.24	.38	.38	.10	.52	.38
										13	.35	.37	.28	.09	.46	.45
		I don't know option	19	.52	.24	.24	.10	.48	.42	10	.43	.37	.20	.12	.60	.28
										13	.44	.40	.17	.11	.57	.32
		Extended instructions	19	.40	.28	.32	.09	.43	.48	10	.57	.29	.14	.10	.49	.41
										13	.24	.38	.38	.07	.37	.56
		Video demonstration	19	.55	.14	.31	.08	.40	.52	10	.31	.25	.44	.08	.42	.50
										13	.42	.26	.32	.08	.40	.52
		Reference handout	19	.33	.37	.30	.08	.42	.50	10	.47	.44	.09	.10	.51	.39
										13	.43	.32	.24	.08	.39	.53
1999	Pozzulo & Lindsay	Simultaneous	NR	.80	.00	.20	.02	.11	.87	13	.65	.11	.24	.08	.38	.54
		Fast Elimination	NR	.48	.03	.48	.01	.05	.94	13	.51	.12	.37	.04	.22	.73
		Slow Elimination	NR	.58	.00	.42	.02	.11	.87	13	.62	.16	.22	.05	.27	.68
2002	Dore	N/A	19	.45	.25	.30	.10	.48	.43	7	.30	.49	.22	.11	.53	.36
										11	.50	.38	.13	.13	.63	.25
										15	.41	.28	.31	.10	.48	.43
2003	Pozzulo & Warren	N/A	20	.68	NR	NR	.01	.06	.93	13	.65	NR	NR	.03	.14	.83
2006	Pozzulo & Balfour	Simultaneous, no change	21	.71	.10	.19	.09	.47	.44	10	.50	.42	.08	.14	.69	.17
		Simultaneous, change	21	.33	.27	.40	.03	.17	.80	10	.21	.58	.21	.08	.42	.50
		Elimination, no change	21	.58	.03	.39	.04	.19	.77	10	.45	.10	.45	.09	.44	.48
		Elimination, change	21	.11	.22	.67	.06	.30	.64	10	.10	.40	.50	.07	.36	.57
2006	Pozzulo & Dempsey (1)	Biased	20	-	-	-	.09	.47	.44	11	-	-	-	.15	.75	.11
2006	Pozzulo & Dempsey (2)	Nonbiased	21	-	-	-	.08	.42	.50	10	-	-	-	.13	.64	.24
		Biased	21	-	-	-	.12	.62	.26	10	-	-	-	.17	.83	.00

2007	Keast et al.	Waiter, unbiased	24	.58	.21	.21	.06	.45	.48	12	.45	.28	.27	.09	.62	.29
		Waiter, biased	24	.55	.32	.13	.08	.55	.37	12	.55	.37	.08	.11	.76	.13
		Thief, unbiased	24	.37	.12	.51	.04	.25	.72	12	.19	.36	.45	.06	.42	.53
		Thief, biased	24	.43	.24	.33	.05	.33	.63	12	.21	.60	.19	.10	.72	.18
2007	Saunders	N/A	NR	.49	.28	.23	.10	.49	.41	4	.46	.34	.20	.14	.70	.16
										10	.48	.33	.20	.08	.42	.50
2009	Ball	High stress	20	.50	.30	.20	.08	.58	.33	14	.46	.38	.15	.09	.61	.31
		Low stress	20	.33	.25	.42	.10	.73	.17	14	.77	.08	.15	.09	.61	.31
2009	Pozzulo et al. (2009) (1)	N/A	22	-	-	-	.08	.39	.53	10	-	-	-	.08	.40	.53
2009	Pozzulo et al. (2009) (2)	N/A	21	.67	NR	NR	.08	.42	.50	10	.50	NR	NR	.14	.69	.17
2012	Pozzulo et al. (2012)	Female Target	21	.46	.00	.54	.07	.21	.72	5	.24	.38	.38	.13	.40	.47
		Male Target	21	.85	.15	.00	.08	.25	.67	5	.21	.45	.34	.14	.43	.43
2012	Havard et al. (2012)	Adult Target	21	.44	.12	.44	.04	.33	.63	7	.30	.63	.08	.09	.71	.20
		Child Target	21	.77	.16	.07	.05	.41	.53	7	.74	.24	.02	.05	.41	.54
2012	Humphries et al. (2012)	Simultaneous	20	.70	.17	.13	.07	.33	.60	6	.60	.33	.07	.11	.53	.37
										10	.73	.20	.07	.07	.33	.60
		Elimination	20	.63	.07	.30	.03	.17	.80	6	.57	.33	.10	.11	.56	.33
										10	.63	.20	.17	.07	.36	.57
		Sequential	20	.83	.03	.13	.04	.22	.73	6	.30	.57	.13	.12	.61	.27
										10	.47	.23	.30	.08	.42	.50
2013	Dehon et al. (2013) (1)	Control task, no delay	24	.82	.00	.18	-	-	-	8	.81	.00	.19	-	-	-
										10	.75	.10	.15	-	-	-
										13	.86	.00	.14	-	-	-
		Control task, then delay	24	.83	.00	.17	-	-	-	8	.45	.25	.30	-	-	-
										10	.45	.14	.41	-	-	-

									13	.52	.14	.33	-	-	-
		Delay, then control task	24	.70	.04	.26	-	-	-	8	.33	.29	.38	-	-	-
									10	.48	.14	.38	-	-	-
									13	.67	.05	.29	-	-	-
		Delay, then description	24	.33	.05	.62	-	-	-	8	.10	.14	.76	-	-	-
									10	.14	.19	.67	-	-	-
									13	.29	.19	.52	-	-	-
		Description, no delay	24	.68	.08	.24	-	-	-	8	.67	.10	.24	-	-	-
									10	.55	.09	.36	-	-	-
									13	.71	.14	.14	-	-	-
		Description, then delay	24	.24	.19	.57	-	-	-	8	.27	.36	.36	-	-	-
									10	.45	.20	.35	-	-	-
									13	.24	.29	.48	-	-	-
2013	Dehon et al. (2013) (2)	Control task	24	.68	.11	.21	.04	.21	.75	7	.75	.10	.15	.08	.38	.55
									11	.75	.10	.15	.03	.13	.85
									13	.55	.35	.10	.05	.25	.70
		Description	24	.29	.29	.43	.07	.33	.60	7	.35	.35	.30	.10	.50	.40
									11	.30	.30	.40	.06	.29	.65
									13	.25	.40	.35	.08	.38	.55
2013	Pozzulo et al. (2013)	Simultaneous	20	.59	.24	.17	.10	.51	.39	17	.69	.19	.13	.09	.44	.47
		Sequential	20	.44	.22	.33	.05	.23	.72	17	.38	.09	.53	.09	.47	.44
		Elimination	20	.47	.09	.44	.04	.18	.78	17	.44	.22	.34	.06	.29	.66
2014	Fitzgerald et al. (2014)	High similarity	20	.74	.15	.10	.07	.34	.59	7	.40	.33	.27	.02	.09	.89
									10	.43	.26	.30	.07	.34	.59
									13	.83	.00	.17	.03	.14	.83

		Low similarity	20	.76	.07	.17	.08	.38	.54	7	.67	.00	.33	.05	.25	.71
										10	.77	.00	.23	.06	.28	.67
										13	.75	.00	.25	.10	.48	.43
2014	Morten (2014)	N/A	21	.57	.13	.29	.07	.37	.56	10	.42	.26	.33	.08	.42	.50

Note. Age = Mean age (if only age range was provided, Age = median of range). NR = Not reported. N/A = Not applicable. FA = False alarm (innocent suspect selection). Innocent suspect selections were estimated by dividing the total false positive rate (for culprit-absent lineups) by the number of lineup members. Clifford (1993) did not report the number of lineup members, so innocent suspect selections were not estimated.

Table A2

Lineup response rates for comparisons between young adults and older adults

			Young Adults							Older Adults						
				Culprit Present			Culprit Absent				Culprit Present			Culprit Absent		
Year	Authors (Experiment)	Condition	Age	Hit	Filler	Reject	FA	Filler	Reject	Age	Hit	Filler	Reject	FA	Filler	Reject
1984	Yarmey et al.	Male criminal	21	.28	.17	.55	.03	.26	.70	71	.20	.41	.39	.07	.57	.36
		Female criminal	21	.22	.14	.64	.02	.15	.83	71	.16	.52	.33	.08	.63	.30
		Male victim	21	.27	.22	.52	.04	.35	.61	71	.09	.39	.52	.08	.65	.27
		Female victim	21	.27	.09	.64	.04	.32	.64	71	.13	.53	.34	.07	.57	.36
1994	Scogin et al.	N/A	21	.29	.14	.57	.06	.30	.64	68	.19	.63	.19	.09	.44	.47
1999	Searcy et al.	Thief lineup	24	.26	.42	.32	.11	.53	.37	70	.19	.63	.19	.13	.67	.19
		Salesman lineup	24	-	-	-	.06	.31	.63	70	-	-	-	.07	.37	.56
		Bystander lineup	24	-	-	-	.06	.29	.66	70	-	-	-	.11	.54	.35
		Talker lineup	24	.68	.05	.26	.07	.35	.58	70	.39	.39	.21	.13	.67	.20
2000	Searcy et al.	Simultaneous, biased	24	-	-	-	.14	.69	.17	69	-	-	-	.15	.74	.11
		Simultaneous, unbiased	24	-	-	-	.12	.61	.27	69	-	-	-	.14	.71	.14
		Sequential, unbiased	24	-	-	-	.06	.31	.63	69	-	-	-	.09	.45	.46
2002	Memon et al.	Mug exposure	22	-	-	-	.07	.34	.59	69	-	-	-	.13	.64	.23
		Mug exposure plus context	22	-	-	-	.05	.25	.70	69	-	-	-	.08	.38	.55
		Control	22	-	-	-	.03	.17	.80	69	-	-	-	.08	.42	.50

2002	Memon & Bartlett	Simultaneous, description	21	.21	.16	.63	-	-	-	70	.35	.25	.40	-	-	-
		Simultaneous, control	21	.13	.27	.60	-	-	-	70	.13	.33	.53	-	-	-
		Sequential, description	21	.38	.43	.19	-	-	-	70	.45	.30	.25	-	-	-
		Sequential, control	21	.13	.40	.47	-	-	-	70	.13	.53	.33	-	-	-
2002	Wright & Stroud (1)	Short delay, young target	22	.47	.40	.13	-	-	-	45	.24	.45	.31	-	-	-
		Long delay, young target	22	.29	.60	.12	-	-	-	45	.20	.54	.27	-	-	-
		Short delay, older target	22	.37	.48	.15	-	-	-	45	.47	.45	.09	-	-	-
		Long delay, older target	22	.21	.62	.17	-	-	-	45	.23	.57	.20	-	-	-
2002	Wright & Stroud (2)	Young target	26	.49	.30	.20	.09	.55	.36	48	.32	.45	.23	.10	.57	.33
		Older target	26	.36	.41	.23	.10	.58	.33	48	.43	.36	.21	.09	.55	.36
2003	Memon, Hope, & Bull	Short exposure	19	.29	.43	.29	.15	.75	.10	68	.35	.45	.20	.13	.67	.20
		Long exposure	19	.95	.05	.00	.07	.34	.59	68	.85	.10	.05	.08	.42	.50
2003	Memon, Bartlett et al.	Short delay, young target	19	.41	.23	.36	.10	.48	.43	72	.45	.41	.14	.13	.66	.21
		Long delay, young target	19	.35	.00	.65	.06	.32	.62	72	.26	.43	.30	.15	.76	.09
		Short delay, old target	19	.41	.41	.18	.10	.52	.38	72	.36	.45	.18	.12	.61	.26
		Long delay, old target	19	.35	.25	.40	.09	.44	.48	72	.04	.65	.30	.12	.58	.30
2003	Memon & Gabbert (a)	Simultaneous	20	.47	.33	.20	.09	.44	.47	69	.48	.29	.23	.15	.75	.10
		Sequential	20	.17	.07	.77	.02	.08	.90	69	.21	.38	.41	.07	.33	.60
2003	Memon & Gabbert (b)	Simultaneous, change	21	.60	.20	.20	-	-	-	69	.30	.50	.20	-	-	-

		Sequential, change	21	.32	.00	.68	-	-	-	69	.15	.55	.30	-	-	-
		Simultaneous, no change	21	.68	.08	.24	-	-	-	69	.25	.40	.35	-	-	-
		Sequential, no change	21	.48	.12	.40	-	-	-	69	.25	.50	.25	-	-	-
2003	Rose et al.	Young target	21	.89	.06	.06	.10	.51	.39	72	.44	.44	.11	.16	.79	.06
		Older target	21	.83	.06	.11	.05	.23	.72	72	.33	.22	.44	.11	.56	.33
2004	Memon et al.	N/A	20	-	-	-	.06	.29	.66	71	-	-	-	.08	.38	.55
2005	Rose et al.	Young target	21	.54	.25	.21	.07	.35	.58	71	.50	.38	.13	.10	.52	.38
		Older target	21	.67	.13	.21	.04	.21	.75	71	.29	.33	.38	.08	.38	.54
2005	Wilcock et al.,	Young target	21	.54	.08	.38	.08	.38	.54	71	.38	.25	.38	.08	.42	.50
		Older target	21	.75	.08	.17	.06	.28	.67	71	.54	.29	.17	.10	.49	.42
2006	Goodsell	Young target, no mugshot	20	.37	.37	.26	-	-	-	70	.21	.32	.47	-	-	-
		Older target, no mugshot	20	.33	.38	.29	-	-	-	70	.27	.46	.27	-	-	-
		Young target, mugshot	20	.12	.64	.24	-	-	-	70	.11	.47	.42	-	-	-
		Older target, mugshot	20	.23	.50	.27	-	-	-	70	.04	.73	.23	-	-	-
2007	Kinlen et al.	Verbalization	20	.10	NR	NR	-	-	-	72	.53	NR	NR	-	-	-
		Visualization	20	.40	NR	NR	-	-	-	72	.53	NR	NR	-	-	-
		Control	20	.35	NR	NR	-	-	-	72	.35	NR	NR	-	-	-
2007	Wilcock et al.,	Young target, context	20	.33	.00	.67	.03	.14	.83	73	.36	.64	.00	.10	.49	.42
		Older target, context	20	.83	.00	.17	.08	.42	.50	73	.45	.36	.18	.08	.42	.50

		Young target, no context	20	.62	.08	.31	.03	.14	.83	73	.33	.42	.25	.13	.63	.25
		Older target, no context	20	.46	.46	.08	.10	.49	.42	73	.67	.25	.08	.15	.76	.08
2009	Havard & Memon	Young target	25	.55	.18	.27	.05	.37	.58	70	.23	.45	.32	.07	.55	.38
		Older target	25	.43	.35	.22	.07	.57	.36	70	.24	.57	.19	.09	.69	.23
2010	Wilcock & Bull	Control	21	.75	.19	.06	.05	.26	.69	69	.63	.19	.19	.13	.63	.25
		Practice	21	.59	.24	.18	.05	.26	.69	69	.56	.19	.25	.06	.31	.63
		Pre-lineup questions	21	.65	.06	.29	.05	.26	.69	69	.38	.44	.19	.06	.31	.63
2010	Houston	Angry mood	20	.29	.48	.24	.11	.57	.32	70	.53	.29	.18	.12	.59	.29
		Neutral mood	20	.52	.29	.19	.11	.56	.33	70	.65	.24	.12	.08	.39	.53
2014	Morten	N/A	21	.57	.13	.29	.07	.37	.56	71	.36	.28	.36	.09	.46	.45
2014	Rochon	Simultaneous	25	.53	NR	NR	.02	.12	.86	77	.29	NR	NR	.17	.83	.00
		Elimination	25	.31	NR	NR	.08	.39	.53	77	.63	NR	NR	.02	.10	.88
n.d.	Key/Gronlund et al.	Fair lineup	26	.69	.11	.20	.09	.46	.45	46	.40	.47	.13	.10	.48	.43
										66	.47	.20	.32	.10	.51	.39
		Biased lineup	26	.83	.03	.14	.09	.46	.44	46	.87	.03	.10	.11	.54	.35
										66	.65	.10	.25	.10	.49	.42
n.d.	Havard	Control	27	.52	.16	.32	.04	.36	.60	77	.48	.36	.16	.08	.68	.24
		Wildcard	27	.56	.24	.20	.05	.39	.56	77	.44	.36	.20	.05	.43	.52

Note. Age = Mean age (if only age range was provided, Age = Median of range). NR = Not reported. N/A = Not applicable. FA = False alarm (innocent suspect selection). Innocent suspect selections were estimated by dividing the total false positive rate (for culprit-absent lineups) by the number of lineup members.

Table A3

Lineup response rates for comparisons between children of difference ages

			Younger Children							Older Children						
				Culprit Present			Culprit Absent				Culprit Present			Culprit Absent		
Year	Authors (Experiment)	Condition	Age	Hit	Filler	Reject	FA	Filler	Reject	Age	Hit	Filler	Reject	FA	Filler	Reject
1979	Marin et al.	Female witness	6	.67	.33	-	-	-	-	9	.50	.50	-	-	-	-
			6	.67	.33	-	-	-	-	13	.83	.17	-	-	-	-
			9	.50	.50	-	-	-	-	13	.83	.17	-	-	-	-
		Male witness	6	.41	.59	-	-	-	-	9	.41	.59	-	-	-	-
			6	.41	.59	-	-	-	-	13	.67	.33	-	-	-	-
			9	.41	.59	-	-	-	-	13	.67	.33	-	-	-	-
1980	Goetze	N/A	8	.42	.38	.21	-	-	-	11	.25	.25	.50	-	-	-
			8	.42	.38	.21	-	-		13	.17	.46	.38	-	-	-
			11	.25	.25	.50	-	-	-	13	.17	.46	.38	-	-	-
1984	King	Live event	7	.46	NR	NR	-	-	-	10	.77	NR	NR	-	-	-
			7	.46	NR	NR	-	-	-	12	.82	NR	NR	-	-	-
			7	.46	NR	NR	-	-	-	17	1.00	NR	NR	-	-	-
			10	.77	NR	NR	-	-	-	12	.82	NR	NR	-	-	-
			10	.77	NR	NR	-	-	-	17	1.00	NR	NR	-	-	-
			12	.82	NR	NR	-	-	-	17	1.00	NR	NR	-	-	-

		Slideshow event	7	.15	NR	NR	-	-	-	10	.25	NR	NR	-	-	-
			7	.15	NR	NR	-	-	-	12	.33	NR	NR	-	-	-
			7	.15	NR	NR	-	-	-	17	.27	NR	NR	-	-	-
			10	.25	NR	NR	-	-	-	12	.33	NR	NR	-	-	-
			10	.25	NR	NR	-	-	-	17	.27	NR	NR	-	-	-
			12	.33	NR	NR	-	-	-	17	.27	NR	NR	-	-	-
1986	Brigham et al.	N/A	10	.68	NR	NR	-	-	-	14	.88	NR	NR	-	-	-
			10	.68	NR	NR	-	-	-	17	.93	NR	NR	-	-	-
			14	.88	NR	NR	-	-	-	17	.93	NR	NR	-	-	-
1986	Goodman & Reed	N/A	4	.38	.31	.31	-	-	-	7	.95	.05	.00	-	-	-
1986	Soppe	Target 1	10	.78	NR	NR	-	-	-	13	.68	NR	NR	-	-	-
		Target 2	10	.33	NR	NR	-	-	-	13	.31	NR	NR	-	-	
1988	Davies et al.	N/A	8	.63	.25	.13	.07	.80	.13	10	.63	.31	.06	.04	.46	.50
			8	.63	.25	.13	.07	.80	.13	12	.69	.19	.13	.05	.52	.44
			10	.63	.31	.06	.04	.46	.50	12	.69	.19	.13	.05	.52	.44
1989	Davies et al.	Mr Nobody	7	.44	.25	.31	.01	.05	.94	11	.69	.06	.25	.02	.11	.88
		Control	7	.50	.13	.38	.05	.33	.63	11	.75	.19	.06	.03	.22	.75
1990	Hammes	N/A	5	.73	.20	.07	.08	.33	.58	11	1.00	.00	.00	.03	.10	.88
1991	dePerczel	Involved female witness	6	.13	NR	NR	-	-	-	12	.38	NR	NR	-	-	-

		Observer female witness	6	.00	NR	NR	-	-	-	12	.25	NR	NR	-	-	
		Involved male witness	6	.25	NR	NR	-	-		12	.38	NR	NR	-	-	-
		Observer male witness	6	.75	NR	NR	-	-	-	12	.13	NR	NR	-	-	-
1991	Goodman Bottoms et al.	N/A		-	-	-	.13	.67	.20		-	-	-	.09	.43	.49
1991	Goodman Hirschman et al.	N/A	4	.30	.60	.10	-	-	-	6	.54	.32	.14	-	-	-
1991	Lieppe	Intruder	6	.39	.23	.39	-	-	-	10	.45	.23	.32	-	-	-
		Toucher	6	.79	.00	.21	.06	.28	.67	10	.63	.19	.19	.02	.11	.87
1991	Oats & Shrimpton	Long delay	5	.33	NR	NR	-	-	-	8	.83	NR	NR	-	-	-
		Short delay	5	.83	NR	NR	-	-	-	8	.91	NR	NR	-	-	-
1996	Clifford & Toplis	Female target	6	.74	.26	.00	-	-	-	9	.22	.17	.61	-	-	-
			6	.74	.26	.00	-	-	-	12	.39	.39	.22	-	-	-
			9	.22	.17	.61	-	-	-	12	.39	.39	.22	-	-	-
		Male target	6	.16	.26	.58	-	-	-	9	.00	.44	.56	-	-	-
			6	.16	.26	.58	-	-	-	12	.11	.28	.61	-	-	-
			9	.00	.44	.56	-	-	-	12	.11	.28	.61	-	-	-
1997	Lindsay et al. (1)	Seq	9	.65	NR	NR	.13	.66	.21	13	.71	NR	NR	.13	.67	.20
		Sim	9	.71	NR	NR	.12	.60	.28	13	.80	NR	NR	.11	.56	.33
1997	Pozzulo & Lindsay	Control	10	.24	.38	.38	.10	.52	.38	13	.35	.37	.28	.09	.46	.45
		Extended instructions	10	.43	.37	.20	.12	.60	.28	13	.44	.40	.17	.11	.57	.32

		I don't know option	10	.57	.29	.14	.10	.49	.41	13	.24	.38	.38	.07	.37	.56
		Reference handout	10	.31	.25	.44	.08	.42	.50	13	.42	.26	.32	.08	.40	.52
		Video demonstration	10	.47	.44	.09	.10	.51	.39	13	.43	.32	.24	.08	.39	.53
2002	Dore	N/A	7	.30	.49	.22	.11	.53	.36	11	.50	.38	.13	.13	.63	.25
			7	.30	.49	.22	.11	.53	.36	15	.41	.28	.31	.10	.48	.43
			11	.50	.38	.13	.13	.63	.25	15	.41	.28	.31	.10	.48	.43
2002	Eisen et al.	Clinician target	4	.78	.22	.00	-	-	-	8	.94	.06	.00	-	-	-
			4	.78	.22	.00	-	-	-	13	.97	.03	.00	-	-	-
			8	.94	.06	.00	-	-	-	13	.97	.03	.00	-	-	-
		Doctor target	4	.68	.29	.03	-	-	-	8	.92	.08	.00	-	-	-
			4	.68	.29	.03	-	-	-	13	1.00	.00	.00	-	-	-
			8	.92	.08	.00	-	-	-	13	1.00	.00	.00	-	-	-
		Nurse target	4	.66	.32	.03	-	-	-	8	.90	.05	.06	-	-	-
			4	.66	.32	.03	-	-	-	13	.89	.06	.05	-	-	-
			8	.90	.05	.05	-	-	-	13	.89	.06	.06	-	-	-
2004	Freire et al.	N/A	4	.30	.32	.38	.10	.48	.42	7	.51	.16	.33	.07	.37	.56
			4	.30	.32	.38	.10	.48	.42	10	.69	.10	.21	.05	.26	.69
			4	.30	.32	.38	.10	.48	.42	13	.72	.08	.21	.03	.16	.80
			7	.51	.16	.33	.07	.37	.56	10	.69	.10	.21	.05	.26	.69

			7	.51	.16	.33	.07	.37	.56	13	.72	.08	.21	.03	.16	.80
			10	.66	.09	.25	.05	.24	.71	13	.72	.08	.21	.03	.16	.80
2004	Huneycutt	N/A	4	.38	NR	NR	.07	.37	.56	7	.39	NR	NR	.13	.67	.20
2005	Brewer & Day	N/A	10	.23	.39	.38	.13	.88	-	16	.44	.14	.42	.13	.88	-
2005	Cain et al.	N/A	4	.37	.63	-	-	-	-	5	.75	.25	-	-	-	-
2006	Beresford & Blades	Elim, photo, SI	7	.43	.14	.43	.05	.38	.57	10	.46	.27	.27	.04	.28	.68
		Elim, video, SI	7	.43	.29	.29	.04	.34	.62	10	.24	.10	.67	.07	.55	.38
		Seq, video, MI	7	.52	.10	.38	.04	.30	.67	10	.38	.38	.24	.06	.47	.48
		Seq, video, SI	7	.48	.30	.22	.08	.67	.25	10	.52	.43	.05	.07	.59	.33
		Sim, photo, MI	7	.43	.29	.29	.05	.42	.52	10	.55	.23	.23	.06	.44	.50
		Sim, photo, SI	7	.57	.19	.24	.07	.59	.33	10	.36	.36	.27	.09	.73	.18
2006	Ross et al.	Bystander lineup	6	-	-	-	.10	.39	.51	8	-	-	-	.15	.61	.24
			6	-	-	-	.10	.39	.51	10	-	-	-	.16	.63	.22
			6	-	-	-	.10	.39	.51	11	-	-	-	.15	.59	.27
			8	-	-	-	.15	.61	.24	10	-	-	-	.16	.63	.22
			8	-	-	-	.15	.61	.24	11	-	-	-	.15	.59	.27
			10	-	-	-	.16	.63	.22	11	-	-	-	.15	.59	.27
		Control lineup	6	-	-	-	.13	.52	.35	8	-	-	-	.12	.46	.42
			6	-	-	-	.13	.52	.35	10	-	-	-	.15	.58	.27

			6	-	-	-	.13	.52	.35	11	-	-	-	.14	.55	.31
			8	-	-	-	.12	.46	.42	10	-	-	-	.15	.58	.27
			8	-	-	-	.12	.46	.42	11	-	-	-	.14	.55	.31
			10	-	-	-	.15	.58	.27	11	-	-	-	.14	.55	.31
2007	Saunders	N/A	4	.46	.34	.20	.14	.70	.16	10	.48	.33	.20	.08	.42	.50
2010	Havard et al.	Photo lineup	8	.45	.28	.28	.07	.57	.36	14	.73	.12	.15	.06	.46	.48
		Video lineup	8	.71	.21	.07	.08	.64	.28	14	.64	.28	.08	.03	.21	.76
2010	Pozzulo et al.	12-member lineup	9	.33	.33	.33	.03	.31	.67	12	.40	.20	.40	.04	.42	.54
		6-member lineup	9	.36	.50	.14	.11	.53	.36	12	.10	.30	.60	.07	.33	.60
2011	Karageorge & Zajac	Control	6	.58	.19	.23	.12	.59	.29	9	.82	.09	.09	.12	.58	.30
		Wildcard	6	.57	.22	.22	.03	.13	.85	9	.71	.19	.10	.03	.14	.83
2012	Clifford et al.	Long delay	8	.41	.24	.35	.09	.71	.21	13	.39	.45	.16	.07	.55	.39
		Short delay	8	.79	.17	.03	.05	.37	.59	13	.50	.33	.17	.07	.56	.37
2012	Humphries et al.	Elim	6	.57	.33	.10	.11	.56	.33	10	.63	.20	.17	.07	.36	.57
		Seq	6	.30	.57	.13	.12	.61	.27	10	.47	.23	.30	.08	.42	.50
		Sim	6	.60	.33	.07	.11	.53	.37	10	.73	.20	.07	.07	.33	.60
2012	Therrien	Control lineup	7	.80	.00	.20	.11	.53	.36	9	.88	.00	.13	.10	.50	.40
		Practice, dog lineup	7	.54	.13	.20	.13	.66	.21	9	.60	.20	.13	.07	.32	.61
		Practice, female lineup	7	.40	.33	.33	.09	.46	.45	9	.50	.14	.20	.10	.48	.43

2013	Havard & Memon	"Mystery Man"	6	.53	.17	.30	.04	.29	.68	10	.57	.36	.07	.05	.43	.51
		Control	6	.46	.39	.15	.08	.64	.29	10	.63	.29	.09	.09	.69	.22
2013	von Zedlitz-Neukirch	Elim	5	.23	NR	NR	-	-	-	10	.34	NR	NR	-	-	-
		Seq	5	.19	NR	NR	-	-	-	10	.28	NR	NR	-	-	-
2014	Dehon et al. (1)	Control task, no delay	8	.81	.00	.19	-	-	-	10	.75	.10	.15	-	-	-
		Control task, no delay	8	.81	.00	.19	-	-	-	13	.86	.00	.14	-	-	-
		Control task, no delay	10	.75	.10	.15	-	-	-	13	.86	.00	.14	-	-	-
		Control task, then delay	8	.45	.25	.30	-	-	-	10	.46	.14	.41	-	-	-
		Control task, then delay	8	.45	.25	.30	-	-	-	13	.52	.14	.33	-	-	-
		Control task, then delay	10	.46	.14	.41	-	-	-	13	.52	.14	.33	-	-	-
		Delay, then control task	8	.33	.29	.38	-	-	-	10	.48	.14	.38	-	-	-
		Delay, then control task	8	.33	.29	.38	-	-	-	13	.67	.05	.29	-	-	-
		Delay, then control task	10	.48	.14	.38	-	-	-	13	.67	.05	.29	-	-	-
		Delay, then description	8	.10	.14	.76	-	-	-	10	.14	.19	.67	-	-	-
		Delay, then description	8	.10	.14	.76	-	-	-	13	.29	.19	.52	-	-	-
		Delay, then description	10	.14	.19	.67	-	-	-	13	.29	.19	.52	-	-	-
		Description, no delay	8	.67	.10	.24	-	-	-	10	.55	.09	.36	-	-	-
		Description, no delay	8	.67	.10	.24	-	-	-	13	.71	.14	.14	-	-	-
		Description, no delay	10	.55	.09	.36	-	-	-	13	.71	.14	.14	-	-	-

		Description, then delay	8	.27	.36	.36	-	-	-	10	.45	.20	.35	-	-	-
		Description, then delay	8	.27	.36	.36	-	-	-	13	.24	.29	.48	-	-	-
		Description, then delay	10	.45	.20	.35	-	-	-	13	.24	.29	.48	-	-	
2014	Dehon et al. (2)	Control task	7	.75	.10	.15	.08	.38	.55	11	.75	.10	.15	.03	.13	.85
		Control task	7	.75	.10	.15	.08	.38	.55	13	.55	.35	.10	.05	.25	.70
		Control task	11	.75	.10	.15	.03	.13	.85	13	.55	.35	.10	.05	.25	.70
		Description	7	.35	.35	.30	.03	.17	.80	11	.30	.30	.40	.06	.29	.65
		Description	7	.35	.35	.30	.03	.17	.80	13	.25	.40	.35	.08	.38	.55
		Description	11	.30	.30	.40	.06	.29	.65	13	.25	.40	.35	.08	.38	.55
2014	Fitzgerald et al. (1)	High similarity	7	.05	.21	.74	.04	.18	.79	10	.13	.50	.38	.08	.38	.55
		High similarity	7	.05	.21	.74	.04	.18	.79	13	.00	.75	.25	.05	.26	.69
		High similarity	10	.13	.50	.38	.08	.38	.55	13	.00	.75	.25	.05	.26	.69
		Low similarity	7	.25	.25	.50	.04	.19	.77	10	.18	.12	.71	.07	.33	.60
		Low similarity	7	.25	.25	.50	.04	.19	.77	13	.25	.13	.63	.07	.35	.58
		Low similarity	10	.18	.12	.71	.07	.33	.60	13	.25	.13	.63	.07	.35	.58
2014	Fitzgerald et al. (2)	High similarity	7	.40	.33	.28	.02	.09	.90	10	.52	.21	.28	.06	.29	.65
		Low similarity	7	.67	.00	.33	.05	.25	.71	10	.77	.00	.23	.07	.34	.59
2014	Rush et al.	High stress, high support	8	.55	.36	.09	.05	.25	.70	13	.64	.00	.36	.05	.25	.70
		High stress, low support	8	.60	.20	.20	.15	.73	.13	13	.40	.30	.30	.07	.33	.60

		Low stress, high support	8	.50	.00	.50	.15	.74	.11	13	.55	.18	.27	.11	.56	.33
		Low stress, low support	8	.56	.11	.33	.11	.53	.36	13	.91	.00	.09	.07	.37	.56
n.d.	Bruer et al.	Categorical ID	7	.09	.27	.64	.06	.44	.50	10	.52	.22	.26	.08	.56	.36
		Categorical ID	7	.09	.27	.64	.06	.44	.50	13	.40	.30	.30	.07	.51	.42
		Categorical ID	10	.52	.22	.26	.08	.56	.36	13	.40	.30	.30	.07	.51	.42
n.d.	Fitzgerald & Price	Target 1, no wildcard	10	.75	.13	.13	.10	.52	.38	13	.58	.17	.25	.07	.33	.60
		Target 1, wildcard	10	.79	.04	.17	.12	.60	.29	13	.77	.00	.23	.08	.42	.50
		Target 2, no wildcard	10	.17	.17	.67	.07	.37	.56	13	.40	.20	.40	.07	.36	.57
		Target 2, wildcard	10	.56	.11	.33	.04	.21	.75	13	.33	.11	.56	.06	.28	.67
n.d.	Fitzgerald et al.	High similarity, target 1	7	.94	.06	.00	.02	.11	.86	10	.75	.08	.17	.07	.34	.59
		High similarity, target 2	7	.60	.10	.30	.07	.32	.61	10	.46	.46	.09	.12	.60	.29
		Low similarity, target 1	7	.80	.08	.12	.05	.25	.70	10	.91	.00	.09	.04	.18	.79
		Low similarity, target 2	7	.73	.18	.09	.10	.48	.42	10	.50	.25	.25	.06	.28	.67
n.d.a	Price & Fitzgerald	Face-off, female target	7	.39	.22	.39	.03	.22	.75	10	.67	.10	.24	.03	.23	.74
		Face-off, male target	7	.50	.05	.45	.08	.58	.33	10	.48	.19	.33	.06	.44	.50
		Sim, female target	7	.35	.24	.41	.07	.49	.44	10	.43	.33	.24	.06	.41	.53
		Sim, male target	7	.56	.28	.16	.03	.21	.77	10	.71	.18	.12	.07	.46	.48
n.d.b	Price & Fitzgerald	High similarity, elim	7	.55	.00	.46	.05	.34	.62	13	.50	.10	.40	.05	.35	.60
		High similarity, elim	7	.55	.00	.46	.05	.34	.62	10	.68	.16	.16	.02	.13	.85

High similarity, elim	10	.68	.16	.16	.02	.13	.85	13	.50	.10	.40	.05	.35	.60
High similarity, face-off	7	.40	.00	.60	.03	.20	.77	10	.50	.10	.40	.02	.17	.81
High similarity, face-off	7	.40	.00	.60	.03	.20	.77	13	.82	.00	.18	.06	.44	.50
High similarity, face-off	10	.50	.10	.40	.02	.17	.81	13	.82	.00	.18	.06	.44	.50
High similarity, sim	7	.67	.33	.00	.03	.22	.75	10	.65	.00	.35	.07	.46	.48
High similarity, sim	7	.67	.33	.00	.03	.22	.75	13	.75	.00	.25	.08	.55	.38
High similarity, sim	10	.65	.00	.35	.07	.46	.48	13	.75	.00	.25	.08	.55	.38
Low similarity, elim	7	.17	.00	.83	.04	.26	.70	10	.62	.10	.29	.02	.15	.83
Low similarity, elim	7	.17	.00	.83	.04	.26	.70	13	.90	.00	.10	.03	.18	.80
Low similarity, elim	10	.62	.10	.29	.02	.15	.83	13	.90	.00	.10	.03	.18	.80
Low similarity, face-off	7	.42	.08	.50	.01	.07	.92	10	.54	.00	.46	.01	.09	.90
Low similarity, face-off	7	.42	.08	.50	.01	.07	.92	13	.64	.00	.36	.04	.29	.67
Low similarity, face-off	10	.54	.00	.46	.01	.09	.90	13	.64	.00	.36	.04	.29	.67
Low similarity, sim	7	.55	.09	.36	.05	.34	.62	10	.67	.08	.25	.05	.38	.57
Low similarity, sim	7	.55	.09	.36	.05	.34	.62	13	.46	.09	.46	.06	.39	.56
Low similarity, sim	10	.67	.08	.25	.05	.38	.57	13	.46	.09	.46	.06	.39	.56

Note. Age = Mean age (if only age range was provided, Age = Median of range). NR = Not reported. N/A = Not applicable. Sim = Simultaneous. Seq = sequential. Elim = elimination. SI = Standard instructions. MI = Modified instructions. FA = False alarm (innocent suspect selection). Innocent suspect selections were estimated by dividing the total false positive rate (for culprit-absent lineups) by the number of lineup members.

Appendix B

Young adults versus older adults: Procedural influences on age differences (within-study effects)

					Weighted Means		Effect Size & CI95			Moderator Test		
					Young	Older						
Procedures Compared	Outcome	Procedure	*m*	*k*	Adults	Adults	ES	*LL*	*UL*	*Q*	*df*	*p*
Lineup vs. Showup	Hits	Lineup	1	2	.75	.56	2.41	1.50	3.86	1.06	1	.304
		Showup	1	1	.53	.39	1.73	1.15	2.61			
	Correct Rejects	Lineup	1	2	.45	.40	1.18	0.81	1.73	0.94	1	.332
		Showup	1	1	.73	.75	1.12	0.73	1.71			
	Sensitivity	Lineup	1	2	2.05	1.48	0.18	0.04	0.32	0.56	1	.455
		Showup	1	1	0.67	0.40	0.11	-0.03	0.24			
	Diagnosticity	Lineup	1	2	8.20	5.81	1.43	0.79	2.63	0.17	1	.681
		Showup	1	1	1.93	1.56	1.24	0.85	1.80			
	Suspect Bias	Lineup	1	2	0.34	0.59	0.16	0.05	0.27	0.05	1	.830
		Showup	1	1	0.26	0.47	0.14	0.06	0.33			
Elimination vs. Simultaneous	Hits	Elimination	1	1	.31	.63	3.66	0.61	21.67	3.07	1	.080
		Simultaneous	1	1	.53	.29	2.89	0.41	19.58			
	Correct Rejects	Elimination	1	1	.53	.88	6.23	0.62	62.26	9.46	1	.002
		Simultaneous	1	1	.86	.00	74.93	3.15	17.81			

	Sensitivity	Elimination	1	1	0.93	2.36	0.35	-0.24	0.94	3.65	1	.056
		Simultaneous	1	1	2.06	0.40	0.50	-0.14	1.13			
	Diagnosticity	Elimination	1	1	3.98	30.00	7.52	0.05	1111.41	1.99	1	.158
		Simultaneous	1	1	22.40	1.71	13.07	0.25	677.58			
	Suspect Bias	Elimination	1	1	0.95	0.86	0.05	-0.54	0.63	0.02	1	.889
		Simultaneous	1	1	0.95	0.76	0.11	-0.52	0.73			
Sequential vs. Simultaneous	Hits	Sequential	3	5	.32	.26	1.34	0.71	2.51	0.14	1	.707
		Simultaneous	3	5	.47	.35	1.65	0.92	2.97			
	Filler Selections	Sequential	3	5	.26	.43	2.83	1.45	5.73	0.64	1	.421
		Simultaneous	3	5	.23	.35	1.93	1.03	3.60			
	Incorrect Rejects	Sequential	3	5	.54	.32	2.54	1.42	4.56	2.98	1	.084
		Simultaneous	3	5	.37	.34	1.11	0.61	2.01			
	Correct Rejects	Sequential	2	2	.68	.49	2.49	1.31	4.72	0.29	1	.593
		Simultaneous	2	2	.39	.12	4.66	1.60	13.60			
	Choosing	Sequential	1	1	0.17	0.49	4.83	2.07	11.29	1.07	1	.300
		Simultaneous	1	1	0.67	0.84	2.54	1.07	6.04			
	Sensitivity	Sequential	1	1	1.16	0.68	0.11	-0.24	0.47	0.01	1	.928
		Simultaneous	1	1	1.26	1.00	0.09	-0.26	0.45			
	Diagnosticity	Sequential	1	1	10.00	3.10	3.22	0.13	82.32	0.14	1	.709

	Simultaneous	1	1	5.25	3.23	1.63	0.36	7.46			
Suspect Bias	Sequential	1	1	1.55	1.16	0.19	-0.17	0.54	0.07	1	.799
	Simultaneous	1	1	0.72	0.54	0.12	-0.23	0.48			

Note. For the five lineup response outcomes (hits, filler selections, incorrect rejects, correct rejects, and choosing), the weighted means are proportions and the effect size is odds ratio. For sensitivity, the weighted means are *d'* scores and the effect size is Hedges' *g*. For diagnosticity, the weighted means are relative risks and the effect size is ratio of relative risks. For suspect bias, the weighted means are $c_{suspect}$ scores and the effect size is Hedges' *g*.