The Jury’s Out: An Analysis of Juror Perception of Child Eyewitnesses and Whether Such Perceptions Are Valid

by

Ilaria Schlitz
Class of 2018

A thesis submitted to the faculty of Wesleyan University in partial fulfillment of the requirements for the Degree of Bachelor of Arts with Departmental Honors in Psychology

Middletown, Connecticut April, 2018
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>3</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>4</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>5</td>
</tr>
<tr>
<td>METHOD</td>
<td>18</td>
</tr>
<tr>
<td>RESULTS</td>
<td>19</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>37</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>63</td>
</tr>
<tr>
<td>SUPPLEMENTAL TABLE 1</td>
<td>70</td>
</tr>
</tbody>
</table>
Acknowledgements

First and foremost I would like to say a giant thank you to my advisor, Jessica Karanian. You made the process of writing a thesis so much more enjoyable than I could ever have imagined. Thank you for going on this adventure with me and for keeping me calm and motivated throughout. I only wish you had arrived at Wesleyan sooner.

I am also grateful to my family for their constant love and support. Thank you for always believing in me, and for encouraging me to pursue my interests. You have been there for me every step of the way, and I could not have made it through this process (or college) without you. Thank you for teaching me to go through life full of curiosity.

Thank you Matt, for all of your words of encouragement. Your late night texts and calls were more helpful than you will ever know. Thank you for making me laugh when I was stressed, and for celebrating with me when I accomplished even the smallest goal.

A big thank-you to my friends for being my support system on campus, for cheering me on, and most importantly for reminding me that there is a world outside of the library.

Lastly, thank you to Professor Bob Steele for graciously agreeing to be my second reader.
Abstract

Despite being largely connected, research into juror perception of child eyewitnesses and child memory currently exist in very separate spheres. In this paper I analyzed 23 experiments related to jurors’ beliefs about child eyewitnesses, and integrated the results with what we know about child memory. At the start of this process I expected to find that jury members vastly underestimate the memory and reliability of children. While my analysis did reveal that children are perceived as less reliable than adults under certain circumstances (e.g. if the child speaks in a powerless style), it also revealed that children are frequently perceived as being equally as reliable as adults, if not more so. This is encouraging given that evidence has shown that children are able to form coherent, detailed memories beginning around 5-years-old. Also of note, my analysis revealed that guilt judgments appear to be made separately from judgments of other dependent measures (e.g. accuracy, honesty, confidence). My original intention was to create a set of recommendations for the court system at the end of this process; however, as I discuss in this paper a number of holes in the research emerged that made this difficult. As a result, I was only able to make a preliminary set of recommendations. These can be found in the conclusion.

Keywords: juror bias, child memory, eyewitness testimony
The Jury’s Out: An Analysis of Juror Perception of Child Eyewitnesses and Whether Such Perceptions Are Valid

The Role of Eyewitness Memory Reports in the Courtroom

Eyewitness testimony plays a crucial role in the outcome of criminal cases, and for good reason. Eyewitnesses provide legal decision makers, such as judges and jury members, with access to information that would otherwise not be available. However, human memory is fallible, and as such while eyewitness accounts are valuable, they can also lead to wrongful convictions. According to the Innocence Project, that misidentification of the accused by eyewitnesses is the leading cause of wrongful conviction (~70% of 354; Innocence Project, 2018). Fortunately, with ever improving technology there are now new and better forms of suspect identification, namely DNA testing. In addition, legal scholars have been increasingly vocal about the unreliability of eyewitness testimony. In response to both developments, 19 states in the United States of America have implemented comprehensive identification reform and multiple jurisdictions have introduced jury warnings with respect to the possible unreliability of eyewitness testimony (Innocence Project, 2018).

While these changes are a step in the right direction, there is still a long way to go. The first step is to improve the procedures surrounding eyewitness identification. According to Garrett (2011), in an analysis of the first 250 exonerations, 88% of the interview procedures were found unreliable and/or suggestive. The second step is to ensure that all jurors understand the strengths, and perhaps more importantly the weaknesses of eyewitness testimony. In a meta-analysis of surveys concerning lay knowledge of eyewitness issues, Desmarais and Read
(2011) found that while jurors have a greater understanding of the issues affecting accurate identification than the researchers originally anticipated, 75% of the 16 eyewitness factors studied continue to be beyond the level of comprehension of the average juror (Desmarais & Read, 2011). More specifically, jurors proved to be less knowledgeable about factors beyond the control of the criminal justice system (e.g. accuracy and confidence of witness, child suggestibility, forgetting curve, post-event information), compared to factors that the criminal justice system can control (e.g. question wording, lineup instructions, mug-shot induced bias). This is not entirely surprising given that the impact of the first set of variables can only be estimated, and as a result there is less information available about them. Of the 16 factors Desmarais and Read (2011) identified, 11 of them are estimator variables. In other words, jurors have a poor understanding of the majority of the factors that influence a witnesses’ reliability, and there is little to nothing the criminal justice system can do to influence the impact of many of them. This is not encouraging given the impact of eyewitness testimonies on jury verdicts.

Of present interest, it appears that such juror perceptions of eyewitness testimonies can be biased in predictable ways. For instance, in cases where the eyewitness is an adult, jurors generally tend to overestimate the reliability of the reports, while the opposite occurs in cases where the eyewitness is a child. It appears that this stems from the assumption that a) children’s memory is less reliable, and b) children are more likely to invent information or be influenced by external factors (e.g., parents, investigators). These assumptions make sense given that we have all witnessed children’s wild imaginations, but is there evidence to support them? At the
same time, we anecdotally know that children are taught from a young age to narrate what happened to them (e.g., show and tell, summer vacation reports, dinner conversations). Thus, asking them to describe what they witnessed is not a novel task per say, but is that experience enough to make them reliable witnesses? In the present paper, I will attempt to answer these questions with scientific evidence. First, I will review the literature on autobiographical memory in children. As will be discussed in the next section, the literature overwhelmingly suggests children may have better memory than we give them credit for.

**Autobiographical Memory in Children**

An important component of memory development, especially as it pertains to eyewitness testimony, is autobiographical memory. Autobiographical memory refers to memories about one’s own life and experiences. These memories include elements of both episodic and semantic memory, and tie basic knowledge of experienced events together with inferences and interpretations of these events (Radvansky, 2011). In this section, I first consider how autobiographical memory changes across development in children (*Development of autobiographical memory*). Then, I discuss the fidelity of such autobiographical memories (*Memory fidelity*).

*Development of autobiographical memory*. Autobiographical memory begins to develop around 2-years-old. This coincides with hippocampal development, an increasing concept of self, and language development – all of which are important factors in memory formation (Goodman, Ogle, McWilliams, Narr, & Paz-Alonso,
Language development is particularly important because it allows children to organize their memories, which in turn solidifies such memories and increases retention (i.e., long-term memory). The importance of language holds true all the way into adulthood given that language skills develop long after the toddler years. For instance, children are able to articulate their thoughts faster and more clearly as they learn to pronounce longer, more complex words, which in turn leads to richer memories (Radvansky, 2011). These results are supported by the fact that children begin to understand the connection between past and present self around 5-years-old. Notably, this is a critical time for language development in children (e.g., learning to read). This subjective consciousness of past and present self then continues to develop into adolescence.

It is important to note that autobiographical memory development in children is not purely accomplished through nature. That is, there is a wealth of evidence that indicates development of autobiographical memory in children can be heavily influenced by one’s environment. For instance, children with highly elaborative mothers (i.e., mothers who spend a large amount of time going over the details of an event with their child) tend to have earlier and richer memories (Fivush, 2011). Specifically, Fivush (2011) found that children with highly elaborative mothers were able to create more detailed and coherent narratives of personal experiences by the end of preschool, as compared to children with less elaborative mothers. Interestingly, maternal reminiscing style is highly predictive of children’s autobiographical memory development, even when controlling for child age, language, temperament, self-concept and attachment status (Fivush, 2011). Of note, mothers who are highly-
elaborative during their child’s preschool years continue to be so as their children get older, which suggests that such influence on memory development may persist throughout development (Fivush, 2011). Why does this elaborative reminiscing style yield such memory benefits? It appears that the reason this reminiscing style is so beneficial is because rather than simply telling their child what happened, elaborative mothers ask open-ended questions that encourage children to recall both factual and qualitative information, and build a narrative around the child’s responses. Research indicates that this learning to create narratives is central to memory development since higher quality narratives lead to higher quality recall (Bauer & Larkina, 2016).

Such findings are important when it comes to child eyewitnesses because they demonstrate that even young children are able to report about personal experiences in a detailed and coherent way. Furthermore, as described by Fivush (2011), multiple studies found that young children (3.5 - to 6-years-old) who were interviewed in an elaborative style following a novel event recalled more information, and reported on the event in a more coherent style (e.g. McGuigan & Salmon, 2004; McGuigan & Salmon, 2005; Tessler & Nelson, 1994). These findings held true both when the child was speaking to a parent and to strangers, and as a result is valuable information for the legal field (for more details, see Fivush, 2011).

While it is encouraging to know that under the right circumstances preschool age children are able to create coherent narratives about past events, it is important to note that life narrative does not truly begin to develop until 8-years-old, and grows in complexity and coherence across adolescence (Fivush, 2011). According to Bohn and Berntsen (2008), children establish a relatively stable life script by around 12-years-
old (Bohn & Berntsen, 2008). However, even then the life script is not yet adult-like in terms of coherence. One possible explanation for the delayed development of life narratives is the onset of childhood amnesia between 5 and 7-years-old (Bauer & Larkina, 2016). That is, at some point between 5 and 7-years-old children forget the majority of their experiences prior to that point. Although the phenomenon of childhood amnesia refers to 5 to 7-year-olds, it is important to remember that children’s ability to remember does not suddenly reach adult levels at 8-years-old. Research suggests that 8-year-olds do begin to display a significant increase in number of memories, but also continue to exhibit a higher rate of forgetting than adults (Baur & Larkina, 2016). This increased rate of forgetting persists past 11-years-old (Baur & Larkina, 2016).

While such findings may not sound promising in terms of child witnesses, it is worth remembering that children do not spontaneously lose the ability to remember during these early years. Instead, they lose early memories over time. Therefore, a child who witnesses a crime at 4-years-old will still be able to report on it months or even a year later, but he or she will not necessarily remember it during their teenage years. According to Amso (2017), the reason for this is that as children grow older their hippocampus and front-parietal regions develop and improve resulting in a greater ability to bind, store, and recall events (Amso, 2017). As such, memories made before these regions have developed are not as permanent or detailed. Another positive note on childhood amnesia is that while this phenomenon clearly affects explicit memory, children do retain many implicit memories from before the onset of childhood amnesia (e.g., learning to ride a bike; Goodman et al., 2013).
Memory fidelity. Beyond knowing that children are able to form and recall memories, it is important to pay attention to the quality of these memories. According to Knutsson and Allwood (2014), multiple studies have shown that when asked open-ended questions, like adults, children recall with 85%-90% accuracy (Knutsson & Allwood, 2014). For instance, Poole and White (1991) assessed the effects of repeated questioning on recall. In their study Poole and White (1991) found that when it came to free recall participants (4-years-old, 6-years-old, 8-years-old, and adult) displayed consistent accuracy regardless of age (Poole & White, 1991). This being said, although the children’s accuracy was equivalent to that of an adult when asked open-ended questions, their level of detail was not. Similarly, Poole and White (1991) also found that children’s free recall is less complete than that of adults, meaning that while the information they recall is correct, they remember fewer details overall.

In addition to a higher rate of forgetting, another possible reason why children leave out information is that they may be unaware of what information is important to share (Knutsson & Allwood, 2014). Similarly, young children lack an understanding of what it means to be coherent. For all of these reasons, children require more open and closed questions to create a useful narrative. The trouble, as Poole and White (1991) highlight, is that children, and especially younger children, are less accurate than adults (who also display lower accuracy when it comes to closed-ended questions) when asked option-posing questions (Poole & White, 1991). Option-posing questions refer to questions where the witness is asked to select a response from two or more pre-established possibilities (e.g. yes or no, blue or green). Even
gentle probing questions that introduce no new information have been found to lead to less reliable testimonies in children (Knutsson, Allwood & Johansson, 2011).

One possible explanation for children’s difficulty with closed-ended questions is that younger children tend to struggle with temporality. As Fivush (2011) explains, preschool aged children will use yesterday to refer to any event in the past, and tomorrow to refer to events in the future, but do not fully understand the meaning of the words, and even 8-year-olds struggle with time if an event occurred more than a few months ago (Fivush, 2011). In a study by Habermas and de Silveira (2008) that looked at the ability of children (8-years-olds), adolescents (12-year-olds and 16-year-olds) and young adults (20-year-olds) to narrate personally significant events and place them on a personal timeline, the researchers found that there is a significant increase in temporal coherence between 8-years-old and 12-years-old (Habermas & de Silveira, 2008).

In addition to being less accurate when answering close-ended questions, children are also more suggestible than adults when it comes to cross-examination, misleading information, co-witness information, and feedback from professionals (Knutsson & Allwood, 2014). This is especially true of younger children. Schaaf, Alexander and Goodman (2008) found that on average it took only three misleading questions for a 3-year-old to assent to a false event, while it took six questions for a 5-year-old to assent (Schaaf, Alexander & Goodman, 2008). One possible explanation for high levels of suggestibility in children is that when presented with an authority figure (such as a parent, police officer, or lawyer) children are less able to discount or refute inaccurate information (Radvansky, 2011). Furthermore, the level of
suggestibility increases as confidence decreases, and children may be less confident when being contradicted by an adult. What is interesting is that research has demonstrated that children were more suggestible when it comes to positive information, than negative, and as a result were more likely to confirm false positive or neutral events than false negative ones (Schaaf et al., 2008). This is especially important to note for eyewitness testimony since, in general, children are asked to discuss negative events in these cases.

While there is evidence to support the claim that children perform at the same level as adults on open-ended questions, and more poorly on closed-ended questions, there is also evidence that suggests that children perform more poorly all around. According to a study by Bauer and Larkina (2016), children actually did not do well when asked open-ended questions, and performed significantly worse than adults. Furthermore, they found that younger children performed worse than older children (Bauer & Larkina, 2016). In their study, Bauer and Larkina (2016) found that when asked open-ended questions one year after an event, adults were 56 times more likely to recall that event than 4-year-olds, 36 times more likely than 6-year-olds, and 16 times more likely to recall that event than 8-year-olds. This being said, as time elapsed the difference decreased, and 8-year-olds recalled at a comparable level to adults two years after the initial event. This is important because it demonstrates how memories weaken regardless of the age of the witness, and also how time levels the field to a certain extent. The other important discovery that Bauer and Larkina (2016) reported was that additional cues and the use of “wh- questions” (i.e. who, when, where) significantly improved recall for young children (4-year-olds recall increased
from 37% to 67%) (Bauer & Larkina, 2016). This is important to consider when interviewing children in legal settings.

In addition to biological maturation, another factor that comes with age is the development of memory strategies. One such strategy is the construction of schemas and scripts. This begins around 3-years-old, and like many other parts of memory continues to develop with age (Radvansky, 2011). Structurally (e.g. generality, linguistic organization, temporal organization) a 3-year-old’s report of events for which they have a script is very similar to that of an older child or even an adult (Hudson, Fivush, Kuebli, 1992). What changes is the level of detail included in the reports. Bohn and Berntsen (2008) suggest that at least when it comes to life scripts, narrative coherence develops throughout adolescence and into young adulthood (Bohn & Berntsen, 2008). The reason the ability to develop scripts is important is because they aid individuals in organizing their thoughts and memories of everyday events. The ability to construct scripts of familiar events is important when it comes to eyewitness testimony, because derivations from the script may stand out and be more memorable (Radvansky, 2011). At the same time, there is a risk that once a child has developed a script they will assume that an event has occurred when it has not, simply because it fits the script of what ordinarily happens.

Although age is a major factor in the quality of an individual’s memories, it is not the only one. Two other important considerations include the severity and emotionality of the event. Although the information is mixed, multiple studies suggest that children are better able to distinguish between memories of real and imagined events when the event in question is serious rather than tedious (Knutsson
& Allwood, 2014). Similarly, it appears that there is less of a difference between children’s and adult’s recall of emotional events, when compared to non-emotional events (Knutsson & Allwood, 2014). For instance, a study by Cordon, Melinder, Goodman, and Edelstein (2013) found that while adults had better recognition of neutral images, there was no significant age difference in memory of aversive images, and especially of images that elicited high or medium arousal during the initial viewing (Cordon, Melinder, Goodman & Edelstein, 2013). Furthermore, like adults, research suggests that children tend to better remember the central details of an emotional event, while the peripheral details are more salient in non-emotional memories. This being said, the details that an eyewitness recalls is modulated by his or her age. For instance, younger children tend to focus on a person’s actions, and exterior details, while adults are more effective in their recall of interior details (Knutsson & Allwood, 2014).

Another important factor is personal relevance. Pathman, Samson, Dugas, Cabeza, and Bauer (2011) found that children’s autobiographical memory is better than their episodic memory (Pathman, Samson, Dugas, Cabeza & Baur, 2011). More importantly, they found that 9-11-year-olds were able to recall autobiographical events at an adult-like level, while 7-9-year-olds were not. Meanwhile, both groups performed more poorly than adults when it comes to episodic memory. In another study that looked at personal narratives, Pathman, Doyum, and Bauer (2013) found that when asked to organize pairs photos they took over the course of four weeks both 8-10-year-olds and adults performed above chance, but adults were more often successful than the children, suggesting that temporal understanding is still
developing during middle childhood (approximately 6 to 11-years-old) (Pathman, Doyum & Bauer, 2013). However, when given a larger number of photos to organize temporally, both adults and children performed poorly. This suggests that while children are not as competent as adults in providing temporal information related to two unrelated events, they are just as capable (or just as incapable) as an adult when it comes to ordering strings of unrelated events.

In addition to those previously mentioned, other contributing factors include the relationship between the child and the interviewer, the warmth of the interviewer, the interviewer’s preconceived beliefs, repeated misleading interviewing, and trauma related psychopathology (Goodman, Ogle, McWilliams, Narr, & Paz-Alonso, 2013). This last factor is particularly relevant when a symptom of pathology is dissociation, as high levels of dissociation have been proven to lead to low levels of accuracy.

Finally, it is important to note that adolescents and young adults experience a “reminiscence bump” beginning at 10-15-years-old up until 30-35-years-old, which is almost the opposite of childhood amnesia. During this period, individuals retain a higher number of memories, and the memories are in greater detail. This phenomenon, perhaps more than any other, is evidence that valuable memories develop years before adulthood officially begins.

**Summary.**

- Children as young as 5-years-old are able to form coherent and detailed memories.
• Evidence suggests that 9 to 11-year-olds recall autobiographical events at an adult-like level.
• Although children’s memories tend to include fewer details, the information they do recall is generally correct.
• Children are more suggestible than adults. This is at least in part due to the power dynamic between children and adults.
• Children are able to distinguish between real and imagined events when the event in question is serious, and children’s memories are particularly strong if the event is emotional or personal.

Present study

When it comes to child eyewitness testimony there are really two issues at hand. The first is that jurors have preconceived notions about the reliability of eyewitnesses in general, and child witnesses specifically, and often these beliefs are not grounded in scientific evidence. The second is that most court systems are not equipped with the relevant information to address these biases. As such the aim of the present study is twofold. The first aim is to determine what beliefs jurors hold about child eyewitnesses, and compare these beliefs with the current research on child memory. From there the second aim is to reconcile these two bodies of information in order to propose a potential path forward that equips jurors with the tools necessary to interpret children’s testimonies in the fairest way possible.
Method

Procedure. Multiple searches on Google Scholar were performed using terms related to jury opinions (‘juror perception’, ‘jury biases’, ‘perception of eyewitnesses’, and
‘juror perception of eyewitnesses’) in order to gain a sense of the available literature on juror perceptions of different populations. Following this initial set of searches, a second series of searches on Google Scholar, PsycINFO, and the Wesleyan library OneSearch database were carried out using terms specific to juror perception of child eyewitnesses (‘juror perception of child eyewitnesses’, ‘juror perception of child witnesses’, ‘jury biases against child witnesses’, ‘jury opinion of child eyewitnesses'). In addition to inputting these words into search engines, relevant empirical articles and review articles were cross-referenced to assure a thorough literature search was conducted. Finally, the studies that appeared in these searches were cross-referenced with those cited in books related to child testimony (e.g. Ceci, Ross & Toglia, 1989; Ceci, Ross & Toglia, 1987). Studies that met all of the criteria for inclusion (see below) were included in a table that was used to identify trends in perception as well as possible explanations for these perceptions. In total, 14 studies were included in the present analysis. From those 14 studies, 23 individual experiments were included.

Inclusion Criteria. Empirical studies published between 1982 and 2017 were included in the literature review provided they met all of the following criteria: were peer-reviewed, studied juror perception and biases, focused on children 18 and younger, the children being evaluated were witnesses not victims, the cases being judged were
not ones of sexual abuse. Of the many studies that the various search engines and the
cross-reference procedures turned up, 14 fit all of the previously mentioned criteria.

*Exclusion Criteria.* Child victims were excluded because in cases involving child
victims, and especially cases of sexual assault, there is a greater risk of outside
influence (e.g. psychiatrists). In addition, the role of trauma in memory production is
also likely to be greater in these cases. Limiting the search to studies involving
witnesses allowed for the focus to remain on the juror’s perceptions of the children,
rather than on the factors that influence a child’s reliability. Other criteria for
exclusion were: studies that focused on the role of figures other than the jurors and
children (e.g. influence of lawyers/ judges, lawyers’ biases) and studies that examined
juror perception overall rather than of children specifically.

**Results**

**Analysis of Experimental Design**

After thorough review of the literature, 23 experiments were identified from
the 14 publications that were included in the present analysis. Most relevant to our
present question was how age of an eyewitness impacts juror perceptions, and
whether such perceptions (those modulated by age) alter juror decision-making in
legal contexts. From the 23 experiments, a number of important factors were
identified as important to consider in terms of being able to answer our present
questions. These factors included age of the eyewitness (*Age*), the variables that were
manipulated (*Independent Variables*), the types of assessments administered
(Dependent Measures), the nature of the crime (Crime Type), and the format in which the eyewitness testimony was presented to the mock jurors (Testimony Modality). These factors and how they varied across experiments are discussed below. All results are summarized in Supplemental Table 1.

Age. There was very little consistency in the ages of both children and adults whose testimonies were included in the various studies. Cumulatively the ages ranged from 3-years-old to 74-years-old, but the ages used in the individual studies were more limited. Furthermore, certain studies listed the ages of the children, but referred to the adults under a single heading. This being said a general trend of age brackets did emerge. Studies tended to look at a combination of the following ages: preschool age children (4-years-old and 5-years-old), 6 to 8-year-olds, 9 to 12-year-olds, college-age students (18 to 22-years-old), and “adults” referring to those older than 21 (although some experiments referred to 20-year-olds as adults as well). Of the 17 experiments that included “adult” witnesses, 13 specified the witness’ age (e.g. 30, 40), including Ross, Dunning, Toglia, & Ceci (1990) who studied juror perception of 74-year-olds. The other four studies simply described them as “adults.” For more details, see Supplemental Table 1.

Independent variables. Age was the most commonly manipulated factor, with 21 experiments (91.3%) manipulating this factor. The other factors that were manipulated were: conflicting witness testimony (same age and different age), credibility of witness (high, moderate, low), type of testimony (real or imagined),
testimony modality, judicial warning about child eye-witnesses (present or absent),
rapport building between child and investigator (included or omitted), identification
of the suspect (positive, negative, foil), role in crime (witness or victim), sex of
witness, competency test, number of errors (0, 3, 6), strength of case, testimony
consistency, and powerfulness of speech. In some cases age was manipulated
alongside one of the other factors in the same experiment, while in others the
manipulations occurred within the same study, but different experiments. For more
details, see Supplemental Table 1.

**Dependent measures.** A variety of dependent measures were used across the 23
experiments included in the present analysis. However, many similarities did exist.
Most commonly assessed was the perceived credibility of the eyewitness as indicated
by mock jury members. Furthermore, a number of studies assessed juror perception of
other eyewitness characteristics including accuracy of testimony, confidence,
trustworthiness, honesty, believability, suggestibility, and consistency. Such ratings
of these characteristics were provided on a likert scale (e.g., 1-5) where higher ratings
indicated stronger perceptions of that attribute. That is, a “5” on suggestibility would
indicate that the juror member perceived that eyewitness as the most suggestible
(except in the case of Ross et al. (1990) where higher scores indicated a more positive
reaction. E.g., less suggestible). In addition to the eyewitness characteristics discussed
above, a number of studies assessed juror member decisions regarding the guilt of the
defendant. Among the 23 experiments included in the present study, six studies
measured degree of guilt using a Likert scale, six experiments asked mock jury
members to make binary verdict judgment (guilty, not guilty), and two experiments asked mock jury members to rate his or her confidence in their verdict on a Likert scale. A full list of dependent measures of interest can be found in Supplemental Table 1.

Furthermore, some studies were designed such that mock jurors had the chance to deliberate with one another. In these studies, mock jurors completed the above-described measures pre-deliberation and post-deliberation. In instances that measures were completed post-deliberation, I marked these measures with “(D)” in Supplemental Table 1. If there is no “(D)” present, then it can be assumed that such measures were administered to an individual juror member.

Crime types. Of the 23 experiments included in this analysis, 15 (65.2%) used witness testimonies related to violent events. These events were a fatal car crash where a pedestrian was killed (n = 3), a murder (n = 2), an abduction (n = 3), a minor assault, a violent drug-deal, a car theft where the car was later used in a robbery-murder, and a robbery-murder (n = 4). Of the eight remaining experiments, six (26.1%) of them used the following events: a non-violent story (n=2), a non-violent theft, a non-violent drug deal (n=2), and a medical injection. The remaining two experiments (8.7%; Leippe & Romanczyk, 1989; Ross et al., 1990) used surveys about hypothetical witnesses to collect data rather than asking about a specific witness testimony. For more details, see Supplemental Table 1.
Testimony modality. The two primary formats for communicating testimonies were transcript and video. 11 experiments (47.8%) used transcripts, and six (26.1%) used video. Additionally, two experiments (8.7%) used narrative summaries of testimonies. Three experiments used a combination of modalities (13%). Luus, Wells and Turtle (1995) used a combination of video and still frames in their second experiment, Nikonova and Ogloff (2005) used a combination of video and summary, and as previously mentioned both Leippe and Romanczyk (1989) and Ross et al.’s (1990) third experiment did not use testimony. For more details, see Supplemental Table 1.

Analysis of Experimental Results

Next, I looked closely at the results of each of the above 23 experiments to identify how age of the eyewitness affected the above-described dependent measures. The results are detailed below for each of the main dependent measures of interest, including juror decision-making, perceived credibility of the eyewitness, perceived confidence of the eyewitness, perceived accuracy of eyewitness testimony, perceived truthfulness/trustworthiness/honesty, perceived suggestibility of the eyewitness, speech style, influence of testimony, perceived consistency of testimony, conflicting reports between eyewitnesses, effects of deliberation, and “other”.

Juror decision-making. Of the six experiments that assessed degree of guilt on a Likert scale after hearing eyewitness testimonies, all (6/6 = 100%) experiments failed

As described above, in six experiments mock jury members instead provided a binary decision regarding the guilt of the defendant (guilty, not guilty). Four of the six (66.7%) experiments similarly found no effect of age (Pozzulo, Lemieux, Wells & McCuaig 2006; Bruer & Pozzulo, 2014, Leippe & Romanczyk, 1989). The other two studies (33.3%) found significant effects of age (Leippe & Romanczyk, 1989; Nikonova & Ogloff, 2005). Specifically, one experiment from Leippe and Romanczyk (1989) revealed that mock juror members were more likely to yield a guilty verdict after hearing testimony from 6-year-olds, as compared to 30-year-olds (Leippe & Romanczyk, 1989). This finding held true even after mock jurors had a chance to deliberate. Alternatively, Nikonova and Ogloff (2005) found a significant effect of age when eyewitness testimonies were presented to mock jurors as a video, as compared to a written summary. When video of the eyewitness testimony was presented, the rate of guilty verdicts was highest for 7-year olds and 10-year-olds, as compared to 23-year-olds. There was no effect of age on guilty verdicts when eyewitness testimonies were presented to mock jurors as a written summary.

Furthermore, Bruer and Pozzulo (2014), Leippe and Romanczyk (1989), Nigro el al., (1989), and Nikonova and Ogloff (2005) found significant effects of other independent factors on juror’s judgments of guilt. Bruer and Pozzulo (2014) found that the number of errors committed in the testimony affected judgments of degree of guilt (Bruer & Pozzulo, 2014). Specifically, a testimony with zero errors was more likely to elicit a guilty rating than a testimony with 6 errors. The three-error
condition was not significantly different than either the zero or six error conditions. Leippe and Romanczyk (1989) found there was a significant increase in the likelihood of a guilty verdict as non-eyewitness evidence strength increased (Leippe & Romanczyk, 1989). They also found a near significant interaction between age and case strength. Their analysis revealed that while age did not reliably influence judgments of guilt when the prosecution’s case, outside of the eyewitness testimony, was weak or ambiguous, there were significantly more guilty judgments when the case was strong and the witness was an adult (rather than a 6-year-old or a 10-year-old). Similarly, Nigro et al.’s (1989) analysis revealed a significant main effect of speech style on assessments of degree of guilt, with higher ratings of guilt being associated with more powerful speech. However, they qualify this finding by exposing a significant interaction between age and speech style (Nigro et al., 1989). Defendants received the highest rating of guilt when an 8-year-old presented their testimony in a powerful speech style, and the lowest rating when the 8-year-old spoke in a powerless style. Meanwhile, speech style had little effect on assessments of guilt when the 25-year-old testified. Both of these patterns persisted after deliberation.

In addition, two of these experiments asked mock jury members to indicate how confident they were in their binary verdict. Neither of these experiments revealed a significant effect of age on verdict confidence (Bruer & Pozzulo, 2014; Pozzulo et al., 2006). Bruer & Pozzulo (2014) also did not find a significant effect of errors on verdict confidence.

Lastly, Nikonova and Ogloff (2005) warned some mock jurors about child eyewitnesses. This warning was based on a set of warnings found in the Canadian
Criminal Jury Instructions, and served to alert the jury to the limitations of children (e.g. limited observational abilities, limited recall, limited comprehension, and underdeveloped moral responsibility). The warning also instructed jurors to assess the child’s evidence relative to their mental development, understanding, and ability to communicate (e.g. young children may leave out details of time and place). They found that warnings about child eyewitnesses led to fewer guilty verdicts when 10-year-olds testified, but warning had no effect on guilty verdicts when 7-year-olds testified (Nikonova & Ogloff, 2005).

Perceived credibility of the eyewitness. A number of the included experiments asked mock jurors to provide a rating to indicate how credible he or she believed the eyewitness testimony to be. Of the fourteen experiments, eight experiments (8/14 = 57.1%) revealed that credibility increased along with the age of the eyewitness (Goodman et al., 1987, Pozzulo et al., 2006; Pozzulo & Dempsey, 2009; Leippe & Romanczyk, 1989; Goodman, Bottoms, Herscovici, & Shaver, 1989). One caveat is that Pozzulo & Dempsey (2009) found a positive correlation between age and perceived credibility only when the individual testifying was a bystander. That is, there was no effect of age on credibility ratings when the person testifying was a victim. Also, it is worth noting that the child victim was perceived as being more credible than the child bystander, but there was no effect of role for the adult.

Interestingly, three experiments (3/14 = 21.4%) revealed the opposite effect, such that younger eyewitnesses received higher credibility ratings than older eyewitnesses (Ross et al., 1990; Nigro et al., 1989; Leippe & Romanczyk, 1989).
Leippe and Romanczyk (1989) suggest that one possible explanation for this discrepancy, at least in the case of their experiments, is testimony modality (Leippe & Romanczyk, 1989). As they explain, in their second and third experiments jurors did not see the child, and so had to base their judgments, at least in part, off of assumptions. Meanwhile, jurors were able to witness the child’s testimony in the fourth experiment, which may have led to stereotype disconfirmation. It is likely that this contrast between expectation and reality is responsible for increased perceived credibility. The “high sincerity” of children may also have become more salient in this condition. Ross et al. (1990) also used a visual form of testimony and as such the same reasoning may apply. While testimony modality cannot be entirely responsible for the discrepancy given that Nigro et al. (1989) presented the testimony as a narrative, and two of the eight “positive” experiments utilized video (Goodman et al., 1987; Goodman et al., 1989) it certainly appears to be an important factor. The other three experiments (23.1%) that examined perceived confidence failed to reveal a significant effect of age on credibility ratings (Ross et al., 1990; Leippe & Romanczyk, 1989; Wells, et al., 1989). There is not enough evidence to explain why these three experiments were unable to find an age effect while the other eleven experiments were.

Lastly, in addition to an age effect, Nigro et al. (1989) found that a powerful speech style led to a higher credibility rating than a powerless speech style for both 8-year-olds and 25-year-olds. This finding held true both before and after deliberation.
Perceived confidence of the eyewitness. Five experiments asked mock jurors to indicate how confident the eyewitness appeared. Three experiments \((3/5 = 60\%)\) found an effect of eyewitness age on perceived confidence (Luus et al., 1995; Ross et al., 1990). In two experiments presented in Ross et al. (1990), both found that 74-year-olds received statistically similar confidence ratings as both 8-year-olds and 21-year olds, but 8-year-olds received higher confidence ratings than 21-year-olds. In Luus et al. (1995), the significant effect of age on confidence ratings was qualified by an interaction with testimony modality (photo + imagined testimony vs. video testimony). Specifically, when eyewitness testimony was presented to mock jurors as a video, as opposed to no video, confidence ratings increased for 8-year-olds and 10-year-olds – however, testimony modality did not affect confidence ratings of adult eyewitnesses. The other two experiments \((2/5 = 40\%)\) revealed no effect of age on confidence ratings (Luus et al., 1995; Wells et al., 1989). However, Luus et al. (1995) did find that credibility had an effect on perception of confidence, with confidence ratings increasing along with the credibility of the testimony. The discrepancy between the results of the two Luus et al. (1995) experiments is interesting because it suggests that testimony modality has an effect on the perceived confidence of children, but not adults. Because neither study directly compares written testimony with video or imagined (mock jurors are shown a still-image of the witness and asked to imagine their testimony), it is hard to determine the exact relationship. However, the fact that written testimony revealed no differences between adults and children, while video and imagined testimony did warrants further investigation.
Perceived accuracy of eyewitness testimony. Nearly half of the experiments asked the mock jury members to indicate how accurate they perceived the eyewitnesses’ testimony to be. Of the ten experiments that measured perceived accuracy, two experiments \((2/10 = 20\%)\) revealed that perceived accuracy increases with age (Goodman et al., 1989; Leippe & Romanczyk, 1989). Worth noting is the fact that Leippe and Romanczyk (1989) only found an effect of age on perceived accuracy of recall. They found no age effect for perceived accuracy of recognition. Interestingly, two experiments \((20\%)\) revealed the opposite effect, and found that perceived accuracy decreases with age (Peterson, 1996; Nikonova & Ogloff, 2005). However, Nikonova’s (2005) results are somewhat nuanced. They found that 7-year-olds are perceived to be as accurate as both 10-year olds and 23-year olds, but 23-year-olds are perceived as less accurate than 10-year-olds. Their analysis also revealed that witnesses’ testimonies were perceived as more accurate when presented as a written summary than as a video. An additional two experiments \((20\%)\) found effects of age that did not fit a linear trend (Ross et al., 1990). In their first experiment, in which they used a video testimony, Ross et al. (1990) found that 21-year-olds were perceived as less accurate than 8-year-olds, who were in turn perceived to be as accurate as 74-year-olds. Meanwhile, their third experiment, which was conducted via a survey instead of presenting a testimony, found that 6-year-old were perceived as less accurate than 8-year-olds, who were in turn perceived to be as accurate as 74-year-olds. 21-year-olds were perceived to be more accurate than all three other ages. Interestingly, Ross et al. (1990) found no effect of age when the testimony was presented as a transcript. Luus et al. (1995) also found no main effect of age;
however, they did find effects of other independent variables in both of their experiments. The transcript condition revealed that as witness credibility increased, so did perceived accuracy. In the second experiment, where they compared type of testimony, they found that viewing a video versus a still frame had no effect on perceived accuracy for 12-year-olds and adults, but videos produced higher perceived accuracy of 8-year-olds. In addition to perceived accuracy, Luus et al. (1995) also examined perceived believability. Their first experiment revealed no effect of age, while the second experiment found that presenting the testimony in video format increased believability ratings for 8 year olds, but had no effect for 12-year-olds or adults.

*Perceived truthfulness, honesty, and trustworthiness.* Six experiments examined perceived sincerity of the eyewitness. Four (4/6 = 66.7%) measured perceived truthfulness (Peterson, 1996; Goodman et al., 1989, Ross et al., 1990), one (1/6 =16.7%) measured perceived honesty, and three (3/6 = 50%) measured perceived trustworthiness (Ross et al., 1990; Nikonova & Ogloff, 2005). Interestingly, Ross et al. (1990) measured both perceived truthfulness and perceived trustworthiness in two of their experiments, although they do not explain the distinction between the two in their article. It is also worth noting that Ross et al. (1990) uses the term truthfulness in their first two experiments, and honesty in their third. This may have to do with the fact that the third experiment was a survey where mock jurors were asked to make judgments without having seen or read the testimony.
Of the six experiments, two (33.3%) revealed that perceived truthfulness/honesty increased with age (Ross et al., 1990; Goodman et al., 1989). The Ross et al. (1990) experiment is interesting because 8-year-olds, 10-year-olds and 21-year-olds were all perceived to be equally honest (in fact, the scores were exactly the same for 8-year-olds and 10-year-olds), and all three age groups were perceived to be less honest than 74-year-olds.

Conversely, two experiments (33.3%) revealed a negative correlation between age and perceived truthfulness/trustworthiness (Peterson, 1996; Nikonova & Ogloff, 2005). Ross et al. (1990) also found a relationship between age and both honesty and trustworthiness, but it is not a linear one. In both experiments 21-year-olds were perceived as both less honest and less trustworthy than 8-year-olds, who were in turn perceived as equally honest and trustworthy as 74-year-olds (Ross et al., 1990).

It is also worth noting that Nikonova and Ogloff (2005) and Peterson (1996) both looked for potential effects of modality and sex respectively, but found no significant effect of either. However, Peterson (1996) did find that mock jurors perceived eyewitnesses as more truthful if he or she had taken a competency test, as compared to when the eyewitness did not take a competency test.

*Perceived suggestibility of the eyewitness.* Three experiments investigated mock jury members’ perception of the suggestibility of the eyewitness. Only one (1/3 = 33.3%) found an effect of age (Ross et al., 1990). Specifically, Ross et al. (1990) found that perceived suggestibility decreases with age, such that 6-year-olds were perceived as the most suggestible, followed by 10-year-olds (Ross et al., 1990). The caveat is that
74-year-olds were perceived to be more suggestible than 21-year-olds. One possible explanation is that the elderly, like children, are often viewed as being less trustworthy eyewitnesses. This is a topic of future investigation.

The two other experiments \( \frac{2}{3} = 66.7\% \) that measured perceived suggestibility found no age effect (Nikonova & Ogloff, 2005; Goodman et al., 1989). The most likely reason for this difference in results is that mock jury members in Ross et al. (1990) were responding to a survey, and so were rating how suggestible they expected a child, adult, or elderly witness to be instead of rating the suggestibility of a witness whose testimony they had either witnessed or read. Also related to testimony modality, Nikonova and Ogloff (2005) found that witnesses in the taped testimony condition were viewed as less suggestible than those in the summary condition. This finding makes sense in relation to many of the previously described results, and especially in light of Leippe and Romanczyk’s (1989) suggestion that viewing a child testify, rather than reading the testimony, results in stereotype disconfirmation.

*Speech style.* Speech style plays an important part in how witnesses are perceived. Three experiments measured forcefulness/powerfulness of speech in some capacity. Two experiments \( \frac{2}{3} = 66.7\% \) examined overall forcefulness of speech (Ross et al., 1990). Of these two experiments, one found no effect of age, and the other found that eight-year-olds were perceived as being more forceful than both young adults (21-year-olds) and older adults (74-year-olds). Once again, it was in the video condition that an effect emerged, further highlighting the importance of being able to observe
the witness. Nigro et al. (1989) did not measure perceived power of speech, but instead included speech style as an independent variable, and measured its effect on both degree of guilt and credibility. They found that credibility increased for both 8-year-olds and 25-year-olds when they spoke in a powerful style. This held true after deliberation as well. In terms of guilt ratings, speech style also had an effect, but it is qualified by an interaction with age. Speech style had no effect on ratings when the 25-year-old testified, but had a significant impact on guilt ratings when the child testified. When the child spoke in a powerful style the defendant received the highest guilt rating of all, but when they spoke in a powerless style they received the lowest rating across ages. This finding held true post-deliberation.

*Influence of testimony.* Many of the dependent measures up until now have looked at judgments of the eyewitness, but three experiments also looked at either how much the testimony influenced the mock jury members’ decision of guilt, or the weight given to the testimony. Two experiments (2/3 = 66.7%) looked at influence of testimony and found no effect of age (Ross et al., 1990). The third experiment (1/3 = 33.3%) was a survey that revealed that mock jurors believed they would give less weight to the testimony of a child (6-years-old or 10-years-old) than to that of a young adult (21-years-old) (Ross et al., 1990). Furthermore, mock jurors viewed older adults’ (74-years-old) testimony more negatively than the younger adults’.

*Perceived consistency of testimony.* Also related to the quality of testimony, three experiments investigated the effect of age on perceived consistency of
testimony. Two experiments found no effect of age on how consistent jurors believed the testimonies to be (Ross et al., 1990; Goodman et al., 1989). In a separate experiment Ross et al. 1990 found that testimonies provided by 74-year olds were perceived to be as consistent as testimonies from both 8-year-olds and 21-year-olds, but 21-year-olds were perceived as being less consistent than 8-year-olds.

*Conflicting reports between eyewitnesses.* While most experiments looked at juror perception of a single eyewitness, Newcombe & Bransgrove (2007) decided to examine mock jury members’ ability to detect accuracy when presented with conflicting testimonies. They did so across two different experiments. In the first experiment both witnesses were the same age, and mock jurors were able to detect accuracy above chance for all ages. In the second experiment witnesses were different ages, and an age effect emerged but only when the youngest witness was correct. In other words, jurors were consistently able to detect accuracy above chance when the older witness was correct, and when the nine-year-old was correct when testifying next to an adult. Where mock jury members struggled was when the four-year-old was correct against a nine-year-old or adult.

*Juror Deliberation.* After being presented with all of the evidence, jurors often deliberate in real-life courtrooms. As a result, three experiments examined the deliberation process in some capacity. One experiment (1/3 = 33.3%) assessed the comments made during the deliberation process (Goodman et al., 1987). This experiment revealed that while the number of positive statements made during the
deliberation process about the eyewitness evidence did not vary across age, the number of negative statements decreased as age of the eyewitness increased. Additionally, the memory and perception abilities of 6-year-olds were questioned significantly more than those of 10-year-olds and 30-year-olds. Finally, in regards to jury discussion on the manipulability of the eyewitness, this topic was never brought up when the case presented a 30-year-old eyewitness, it came up occasionally when the case presented a 10-year-old eyewitness, and the topic was of great concern when the case presented a 6-year-old eyewitness.

The remaining two studies (2/3 = 66.7%) compared scores on the same measures before and after deliberation (Leippe & Romanczyk, 1989; Nigro et al., 1989). Nigro et al. (1989) found no change in age effects post-deliberation for either degree of guilt or credibility. With regards to Leippe & Romanczyk (1989) general trends for guilty judgments and credibility with respect to age did not change; however, guilty judgments were less frequent post-deliberation, and overall credibility decreased as well. This being said, Leippe and Romanczyk (1989) do note that the exonerating effect of deliberation is weaker in judgments of 6-year-olds. They also found a significant interaction between age and timing (pre-or post-) for credibility, whereby credibility judgments decreased during deliberation if the witness was an adult, but had no effect on judgments of children.

Other. In addition to the results discussed above, four experiments examined other dependent variables related to juror perception of eyewitnesses. These are: effect of age and rapport building on judgments of child characteristics, interviewer behavior,
child understanding, and case progression (Krähenbühl, 2012), perceived integrity (Brue\& Pozzulo, 2014), and perceived strength of the prosecution’s evidence (Leippe & Romanczyk, 1989).

With regards to the Krähenbühl (2012) experiment, there was no effect of age on child characteristics, interviewer behavior, child understanding, and case progression. However, ratings for all of these variables were higher when mock jurors were shown the rapport-building phase of the interview, compared to when they were not.

Integrity, as it is measured in Bruer & Pozzulo (2014) is actually the result of collapsing the scores for perceived accuracy, perceived credibility, and perceived overall reliability (which was determined by collapsing 3 individual reliability scores). This experiment revealed that adolescents (12-years-old) and adults (20-years-old) are viewed as having significantly more integrity than children (4-years-old). The experiment also revealed an effect of number of errors. Witnesses who committed six errors during their testimony were viewed as having significantly less integrity than those who committed zero or three errors.

Leippe and Romanczyk (1989) found no effect of age on perceived strength of the prosecution’s evidence in their second experiment. However, they did find that the evidence was perceived as being stronger when the witness was consistent. Their third experiment revealed an effect of age, whereby the prosecution’s evidence was perceived as being stronger when the 6-year-old testified than when the 30-year-old testified.
Discussion

In the present paper, I hoped to achieve two goals. The first was to identify the current body of research related to juror perception of child eyewitnesses and compare it to what we know about child memory in order to verify whether jurors’ beliefs about child eyewitnesses are supported by scientific evidence. The second goal was to use this newly compiled information to develop a set of recommendations that would be useful to jury members and other members of the legal profession. In total, 23 experiments from 14 different publications were included in this analysis. A number of interesting findings arose from this process, some of which confirmed my belief that jurors, and laypeople in general, underestimate the memory of children. However, a surprising number of experiments revealed that either there was no effect of age, or that jury members may favor children over adults in certain situations.

Present findings

Juror decision-making. One important finding that came out of this process is that judgments of guilt appear to be independent from judgments of other dependent measures (e.g. accuracy, honesty, consistency, etc.). In other words, the majority of the experiments that measured guilt (10/12 = 83.3%) found no main effect of age on judgments of guilt even when the mock jurors favored one demographic. For instance, Goodman et al. (1987) found no effect of age on judgments of degree of guilt even though mock jury members found 6-year-olds less credible than 10-year-olds and adults, expressed greater concern about the manipulability of 6-year-olds,
expressed greater doubt about young children’s memory and processing abilities, and
overall made more negative statements about their testimony.

One possible explanation for this finding, at least for the experiments where
jurors made judgments about degree of guilt (instead of binary decisions), is that the
majority of the crimes detailed in these experiments have very clear outcomes. In the
murder cases the victim definitively died, in the fatal car crashes the pedestrian was
clearly hit by the car, the children in the abduction scenario were undeniably
abducted, and so on. As a result, the degree to which the witness is perceived as
confident, consistent, trustworthy, etc. does not change how guilty the defendant is.
They either committed the crime or they did not.

Although most experiments were unable to find an effect of age on judgments
of guilt, several experiments did reveal effects of other measures that are worth
discussing. The first is that, not surprisingly, the number of errors a witness made was
negatively correlated with degree of guilt (Bruer & Pozzulo, 2014). Also, Nikonova
and Ogloff (2005) found that testimony modality had an impact on judgments of
guilt. Specifically, when the testimony was presented as a video, the number of guilty
verdicts was higher for 7-year-old and 10-year-old eyewitnesses than for 21-year-old
eyewitnesses. Meanwhile, there was no effect of age if the testimony was presented in
written form. This suggests that being able to hear and see a child testify, rather than
simply reading their testimony, makes the child’s strengths more salient, and disrupts
negative stereotypes about their abilities. As a result, their testimony may be weighted
more heavily in the jurors’ judgments of guilt than it might have otherwise been.
Another factor that interacts with age when it comes to guilty judgments is power of speech. Nigro et al. (1989) found that when a child testified using powerful speech the defendant received the highest rating of guilt out of all the conditions, while when a child spoke in a powerless style, the defendant received the lowest rating of guilt. Speech style had no effect on guilty judgments when the witness was an adult. One potential explanation for this pattern is that adults are already expected to be reliable, and so their speech style has little effect. On the other hand, children are often seen as less reliable, and so if they speak in a powerful manner it significantly boosts the perceived value of their testimony, while if they speak in a powerless way it confirms their unreliability.

Leippe and Romanczyk (1989) also found a relationship between age and another dependent measure. In their case, they found that case strength affected guilty verdicts, but only if the witness was an adult. If the prosecution’s evidence, outside of eyewitness testimony, was weak or ambiguous there was no effect on the verdict, but if the case was strong it increased the number of guilty verdicts for 30-year-olds, but not for 6-year-olds or 10-year-olds. Assuming that the jury heavily weighs an adult’s testimony, then their testimony paired with an already strong case against the defendant is likely to yield a guilty verdict. Whereas if the jury views the child’s testimony with skepticism, then they are less likely to reach a guilty verdict, even if the rest of the case is strong.

A third interaction of age is with pre-trial warnings about the strengths and limitations of child witnesses. The presence of a pre-trial warning resulted in fewer guilty verdicts, but only when the older child (10-years-old) testified (Nikonova &
Ogloff, 2005). It is possible that the shortcomings of child witnesses were already salient to jurors when the 7-year-old testified and so the warning had no effect. In contrast, the warning may have reinforced negative biases about 10-year-old witnesses that the jurors may have otherwise overlooked.

**Perceived credibility of the eyewitness.** Another area in which testimony modality played an important role is in perceived credibility. Of the eight experiments that found a positive correlation between age and perceived credibility, only two (2/8 = 25%) used video testimonies, while two of the three experiments (2/3 = 66.7%) that found a negative correlation used video. What this suggests is that not only does viewing the child’s testimony increase their perceived credibility, but also it does so to such a strong degree that their positive attributes outshine their negative ones, and they are viewed more positively than the adults. Leippe and Romanczyk (1989) posit that the “high sincerity” of children may also become apparent in the video modality.

Three other experiments that measured perceived credibility found no effect of age, but Nigro et al. (1989) did find that powerful speech increases credibility for both children and adults. This finding is not surprising given that powerful speech creates a sense of confidence, and individuals who are confident also appear more credible. This is evidenced by Luus et al. (1995) who found that by manipulating credibility they were also able to control perceived confidence. Specifically, as credibility increased so did ratings of confidence.
Perceived confidence of the eyewitness. A second experiment by Luus et al. (1995), that also measured the perceived confidence of the witness, found a significant effect of age on confidence. This effect, like many others was qualified by an interaction with modality. What the experiment revealed is that when mock jurors were presented with a video testimony, as opposed to being presented with a still image and asked to imagine the testimony, confidence ratings increased for 8-year-olds and 10-year-olds, but not for adults. In a third experiment, where the testimony was presented as a transcript, Luus et al. (1995) found no effect of age. The discrepancy between these two studies is not surprising given what has already been discussed about how video testimonies create a more positive opinion of child witnesses. The fact that Luus et al. (1995) compares three modalities across two studies, and video has the strongest positive effect out of all of them is further evidence of the importance of allowing jurors to experience the child’s testimony first hand.

Also related to confidence, two experiments by Ross et al. (1990) revealed that statistically speaking, 74-year-olds were perceived as being as confident as both 8-year-olds and 21-year olds, but 21-year-olds were perceived as being less confident than 8-year-olds. These findings held true both when the testimony was presented in a video format and in written form, meaning that modality did not have an effect in this case. What these findings tell us is: first, under certain circumstances children may be perceived as more confident than adults. Second, in terms of how they are perceived, elderly adults are perceived as being somewhere between children and adults. This makes sense if one considers the fact that the elderly are believed to experience some of the same shortcomings as children in terms of memory.
Perceived accuracy of eyewitness testimony. The results that emerged about perceived accuracy are very mixed. A little more than half of the experiments that looked at accuracy found age effects (6/10 = 60%). One third (2/6 = 33.3%) found a positive correlation between age and perceived accuracy, one third (2/6 = 33.3%) found a negative correlation, and one third (2/6 = 33/3%) found a non-linear effect of age. Of note is the fact that two of the experiments used surveys rather than testimonies to obtain mock jurors opinions. One survey revealed that without an explicit testimony to guide them, mock jury members believed that accuracy increases with age (Leippe and Romanczyk, 1989). The other survey found 6-year-olds are perceived as being the least accurate, and 21-year-olds are perceived as being the most accurate, but 8-year-olds and 74-year-olds are perceived as being equally accurate (Ross et al., 1990). This second result is interesting because it once again points to the fact that elderly adults are perceived as occupying a middle ground when it comes to memory.

To complicate things further, Nikonova and Ogloff (2005) found that eyewitness testimonies were perceived as more accurate when they were a written summary than when they were presented as a video. This finding contradicts the majority of evidence described up until this point, which has demonstrated that video testimony elicits a more positive opinion of witnesses, and particularly children. Interestingly, Ross et al. (1990) found no effect of age when the testimony was provided in written form, which matches prior observations about testimony modality.
Luus et al. (1995) also found effects of testimony modality. In the written condition they found that accuracy increased along with credibility. In the video vs. still image condition there was no effect of modality on 12-year-olds and adults, but perceived accuracy increased for 8-year-olds in the video condition. The same pattern emerged for believability. The fact that modality did not affect mock jury members’ perception of the adults’ accuracy does not come as a surprise since it has failed to do so in the other experiments as well. What is surprising is the distinction between 8-year-olds and 10-year-olds. What this difference suggests is that 10-year-olds are perceived as being more adult like in their accuracy, and so there is less of an effect of stereotype disconfirmation.

While a large number of experiments measured how accurate mock jury members perceived witnesses to be, only one experiment explored their ability to appropriately detect accuracy. Newcombe and Bransgrove (2007) presented mock jurors with two conflicting testimonies, either from eyewitnesses of the same age, or of different ages. They found that mock jury members were able to detect accuracy at a rate above chance when the two witnesses were the same age, if the older witness was correct, or if the mid-aged child (9-years-old) was correct. Where mock jurors had difficult was in identifying the accurate testimony when the 4-year-old was telling the truth. These results demonstrate that while jurors are capable of detecting accuracy, this ability may be overshadowed by negative biases they hold about young children.
Perceived truthfulness, honesty, and trustworthiness. Like perceived accuracy, the results of the sincerity measures (honesty, truthfulness, and trustworthiness) were quite varied. Of the seven experiments that looked at these measures, two (2/6 = 33.3\%) found that perceived sincerity increased with age, two experiments (2/6 = 33.3\%) found a negative correlation, and three experiments (2/6 = 33.3\%) found a non-linear pattern. Interestingly, in these last two experiments the 21-year-old was perceived as the least honest/trustworthy. This may have had to do with the nature of the crime. In both experiments the witness was testifying about a non-violent drug deal, and so it makes sense that the 21-year-old would be viewed as less honest/trustworthy, given that of all the ages they are the ones most likely to be involved with drugs. The other experiments that looked at sincerity measures did not involve crimes where age would be an important factor.

In addition to an effect of age, Peterson (1996) also found that witnesses who had taken a competency test were perceived as more truthful. This is not surprising since a competency test can be viewed as an extra check of a witnesses’ reliability.

Perceived suggestibility of the eyewitness. Of the four experiments that investigated perceived suggestibility, only one found an effect of age, and that result was based on a survey, not on an experienced testimony (Ross et al., 1990). In the survey, mock jury members stated that they believe suggestibility decreases with age. The caveat is that 74-year-olds were perceived as being more suggestible than 21-year-olds. This result coincides with the other findings about elderly adults. Also not surprising is the fact that Nikonova and Ogloff (2005) found that witnesses in the video condition
were viewed as less suggestible than those who participated in the summary condition. This supports the previously stated belief that viewing testimonies rather than reading them reduces biases.

*Other perceptions of eyewitnesses.* Experiments also measured other types of juror perceptions, including forcefulness, influence on decision-making, consistency, and integrity. However, such results should be interpreted with caution due to such small sample sizes.

Two experiments looked how forceful mock jury members perceived the witnesses to be (Ross et al., 1990). Only the experiment with the video modality found an effect of age, which is that 8-year-olds were viewed as more forceful. One potential explanation for this finding is that mock jurors were not expecting a child to be forceful, and so when they witnessed the child being forceful it stood out more than an adult being forceful. Relatedly, as discussed above, powerful speech was associated with greater credibility for both children and adults, and higher ratings of guilt for children, but not for adults.

Only three experiments investigated perceived influence of the eyewitness’ testimonies. Of the three experiments, two (2/3 = 66.7%) failed to find an effect of age (Ross et al., 1990. The third experiment found that jurors believed that they would give more weight to eyewitness testimonies as the age of the witness increased, the exception again being 74-year-olds who fell in the middle (Ross et al., 1990). Important to note is the fact that this third experiment was a survey, and as a result
revealed a disconnect between how mock jurors believe they would respond and how they actually respond when presented with a child eyewitness.

Of the three experiments that looked at the effect of age on perceived consistency, only one found an effect of age (Ross et al., 1990). Each of the three experiments represented a different testimony modality (including survey), and the video modality was the only one to reveal an effect of age. In this experiment 74-year-olds were perceived as being as consistent as both 8-year-olds and 21-year-olds, but the 21-year-old was perceived as being less consistent than the 8-year-old. Given that only one experiment found an effect of age, it is difficult to know why this pattern emerged. Of relevance, Leippe & Romanczyk (1989) found that the prosecution’s evidence was judged as being stronger when the witness’ testimony was consistent. Additionally, while they found no effect of age on strength of evidence in their third experiment, they did find an effect in their fourth experiment where the strength of the prosecution’s evidence was judged as being stronger when the 6-year-old testified. The reason for this distinction is that although the crime was the same in the two experiments, the latter experiment included more details about the child’s character. These details illustrated the ways in which the child did not fit the stereotypes jurors may believe about children. This once again led to stereotype disconfirmation.

It is not surprising that 12-year-olds and 20-year-olds were seen as having more integrity than 4-year-olds (Bruer & Pozzulo, 2014). Although this analysis has proven that mock jury members hold children in higher regard than I anticipated, it is clear that they still hold certain biases against them. It is also no surprise that
witnesses of any age who committed six errors were viewed as having less integrity than those who committed zero or three errors.

*Effects of juror deliberation.* One of the most interesting findings was about the effect of deliberation. Three experiments studied deliberation in two different ways. Goodman et al. (1987) studied the statements that mock jury members made during the deliberation process. This experiment revealed that the number of negative comments made about the witness decreased as age increased. In addition, the memory and perception of the 6-year-old was questioned significantly more than when the witness was 10-years-old or 30-years-old. The topic of manipulability was never brought up when the witness was 30-years-old, was brought up occasionally when they were 10-years-old, and was a topic of concern when they were 6-years-old. Based on these results it is clear that mock jury members continue to hold negative biases about children, even after they have witnessed their testimony. In addition, it appears that while 10-year-olds are not viewed as equals to adults, they are viewed as being somewhat similar.

Interestingly, Nigro et al. (1989) found no effect of deliberation on either guilt verdicts or credibility. Leippe and Romanczyk (1989) on the other hand found that while the pattern of results did not change, the number of guilty judgments declined post-deliberation, especially if the witness was an adult. Overall credibility also decreased if the witness was an adult. Based on these two studies it appears that deliberation has little to no effect on perception of child-witnesses. Given that many of the previously discussed results point to the fact that viewing a child testify
dismantles certain negative biases, it may be that jurors have already been forced to reconsider their views on child testimony, and so deliberation has less of an effect on their perception. Meanwhile, prior to deliberation jurors have less of a reason to re-think their views on adults testifying.

*Rapport building*: A final interesting finding that arose during this process is that allowing mock jurors to see the rapport building stage of the interview increased ratings of child characteristics, interviewer behavior, child understanding, and case progression (Krähenbühl, 2012). I already knew that rapport-building increases trust between the child and the interviewer, but it was interesting to learn that it making jurors privy to this interaction also increases their positive view of the child and the case in general.

**Integration of results with what we know about childhood memory.**

After thoroughly analyzing the existing literature on juror perception of child eyewitnesses, it has become clear that jurors do in fact hold many of the biases that I expected. At the same time, it has also become apparent that under certain circumstances, especially when they experience stereotype disconfirmation, jury members may in fact perceive children to be equally as reliable as adults, if not more so. This being said, it is not enough to know what jury members think. It is important to take this newfound knowledge, and examine it in the context of research on child memory. Doing so will allow us to determine how much of what jury members believe is grounded in scientific evidence.
As we saw in the introduction, language development is an important part of memory formation since it allows us to organize our memories in a coherent way, which in turn leads to improved retention. From Nigro et al. (1989), we have evidence to suggest that language is also important in juror perception, but for a different reason. Jurors value language abilities more for the confidence and coherence that they provide, than for the effect that articulation has on memory organization and retention, but ultimately the two are highly interconnected. After all, the ability to articulate thoughts quickly and clearly, as well as the ability to pronounce longer and more complicated words are both important in memory development, and these same skills contribute to a powerful speech style.

In addition to power of speech, consistency is another important element of eyewitness testimony. Two out of the three experiments included in this analysis found no effect of age on consistency (Ross et al., 1990; Goodman et al., 1989). This implies that children are perceived as being as consistent as adults in their testimony. According to Fivush (2011), children as young as 5-years-old can develop coherent and detailed narratives. Additionally, children begin developing a sense of past and present around 5-years-old, which in turn contributes to consistency. This being the case, it is entirely plausible that the children involved in the studies mentioned above were in fact consistent in their narratives, and it is encouraging to know that jury members are able to recognize this consistency. Fivush (2011) attributes the ability to develop strong narratives at a young age to maternal reminiscing style, but it is worth considering that as schools shift from rote learning to a more discussion based curriculum, children may be developing memory skills as school as well.
The one study that did find an effect of age on consistency revealed that adults are viewed as less consistent than children when testifying in court (Ross et al., 1990). While there is no experimental evidence to support this theory, it is possible that the reason adults are perceived as being less consistent is because they have moved away from the period in their lives when people are guiding them to develop a linear narrative. As a result they may have lost this habit.

Another aspect of Fivush’s (2011) research that is relevant to the current analysis is the effect of an elaborative interviewing style. As Fivush (2011) explains, children interviewed in an elaborative style recall more information and are more coherent. Not only does this support the possibility that adults other than the mother can influence children’s memory development, it also offers support for the importance of witnessing a child’s testimony. As I determined based on a large number of the studies, video testimony elicits the most positive judgments of child witnesses as compared to other less personal modalities (e.g., summary narrative). Taking Fivush’s (2011) research into account, one possible reason why video (or even better, in person testimony) has such a positive effect on jurors’ perception of children is that it allows them to witness the elaborative questioning style in progress, and by extension observe how coherent children interviewed in this way are capable of being. A child that is coherent and includes plenty of details in their testimony also exudes an air of confidence, which as I saw from my analysis is important when it comes to the jury’s perception of a witness.

A surprising result that emerged during my analysis is that jury members in multiple studies perceived children as being more accurate than adults. This includes
Peterson (1996) whose experiment revealed that mock jury members perceived 7-year-olds as being more accurate than 20-year-olds. Similarly, Nikonova and Ogloff (2005) found that 10-year-olds were believed to be more accurate than adults. Interestingly, in this case, jury members may have been over-estimating children’s abilities. According to Bohn and Berntsen (2008), children develop a stable narrative around age 12, and even then the level of coherence is not equal to that of an adult. Of course, in court children are not being asked to recount their entire life story, but it is still important to consider the fact that their ability to construct a narrative is not fully developed at such a young age. Goodman et al. (1989) found that when comparing children ages 3-years-old to 6-years-old accuracy is believed to increase with age. Given what we know about children’s memory this is an accurate assessment – especially since children do not begin to develop a life narrative until the later preschool years.

What is also interesting about the research on perceived accuracy is that the two surveys (Ross et al., 1990; Leippe & Romanczyk, 1989) reveal a different trend from the experiments mentioned above. In both surveys jurors expressed that they expect accuracy to increase with age, which may be closer to what actually happens. One possible reason for the difference in responses to surveys vs. testimonies is that witnessing a child testify may disrupt jury members’ expectations, and cause them to give more power to the child’s testimony than its contents actually merit. It is also possible that if the child is the only eyewitness jurors may not have any other testimony to compare theirs with, and so it may not be apparent that their narrative is less detailed or coherent than an older child’s testimony or an adult’s testimony.
Although a number of studies found that jury members perceived children to be more accurate than adults, there are also quite few studies that found that children were viewed more negatively, and whose results reflect the fact that life narrative develops with age. For instance, Bruer & Pozzulo (2014) collapsed credibility, accuracy, and reliability to form an integrity rating, and found that integrity scores increased with age. Similarly, both Pozzulo et al. (2006) and Pozzulo and Dempsey (2009) found that 10-year-olds were perceived to be less credible than adults. Not only do the results of these three experiments correspond with the finding that the coherence and level of detail associated with life narrative increase with age, they are also supported by Bauer and Larkin’s (2016) findings that children are less accurate than adults regardless of questioning type.

In addition to the experiments that found that jurors favored one demographic over another, there are a number of experiments that found no effect of age on different dependent measures such as believability, accuracy, and credibility (e.g. Luus et al., 1995; Goodman et al., 1987; Ross et al., 1990). These studies are the ones that are perhaps best supported by Knutsson and Allwood’s (2014) claim that with regards to open-ended questioning children recall at a similar rate to adults. Most of these studies do not make reference to the type of questioning that takes place, nor completeness of testimony, and so it is hard to make a direct comparison – however, based on what we do know, they seem to serve as evidence that at least under some circumstances children are as accurate as adults.

Nevertheless, the conflicting evidence about the effects of questioning style make it difficult to reach any sort of conclusive judgment about the accuracy, credibility,
believability, etc. of child witnesses. As a result I am unable to say whether the jurors in these different experiments are correct in their assessments.

Sumner-Armstrong and Newcombe (2007) do present evidence that suggests juror education can play a significant role in accurate assessments. In their study, Sumner-Armstrong and Newcombe (2007) assessed the effects of a neutral form of jury education (as opposed to an expert witness). The education took the form of a five-page booklet that contained a variety of information including research on child witnesses, information on child memory, and a discussion on interviewing techniques and potential risks. What they found was that jurors who received the information prior to the trial were able to correctly rate the child witness’ accuracy, honesty, reliability, and credibility, whereas those who had not received the education performed at a level no greater than chance. The measures the jurors were tested on are many of the same measures that the experiments in this analysis looked at, and so it is promising to know that with education jurors are able to accurately assess these measures. This study also found that education might be effective in preventing stereotypical assessments.

Another surprising finding is the fact that of the five experiments that looked at perceived suggestibility, only one found an effect of age, and that was the survey, which found jurors expect suggestibility to decrease with age (Ross et al., 1990). Given that none of the other experiments found an effect of age, it appears as though when jurors observe or read a child’s testimony their belief about suggestibility diminishes. Meanwhile there is scientific evidence that shows that there is actually some validity to the assumption that children are more suggestible. For instance, there
is evidence that children are more suggestible than adults because children have a more difficult time discounting or refuting false information when confronted with an authority figure. It is possible that since many of the experiments focus on testimony rather than questioning this power dynamic was less salient, thereby reducing the suggestibility ratings. This being said, the study that looked at rapport did find improved perception all around when rapport was shown, which suggests that jurors are aware of the importance of the interviewer-eyewitness relationship.

Although suggestibility was not a major concern in the studies included in this analysis, it is worth noting that were are ways to mitigate the effects of power dynamics. According to Radvansky (2011), one such way is through repeated interviewing, since this process gives the child the opportunity to work past the initial fear of contradiction.

Suggestibility is also related to confidence, and while none of the experiments that look at suggestibility looked at confidence, in general it appears that the children involved in the studies included in this analysis were perceived as being quite confident, which may have skewed the results. Another reason why the child eyewitnesses may have appeared less suggestible than I originally anticipated is that children are more suggestible when it comes to positive information than negative (Schaaf et al., 2008), and eyewitness testimony generally involves negative information. The exception to this trend of low perceived suggestibility is Goodman et al. (1987) who found that manipulability was a major concern only when the witness was a child.
Although there is evidence that the severity and emotionality of an event is predictive of how accurately a child remembers an event, and what details they remember (Knutsson & Allwood, 2014), there is no clear pattern in how crime type affects juror perception of children across the experiments included in this review.

Finally, research also suggests that children are better at remembering events of personal relevance; however, because the individuals testifying in all of these studies were intentionally witnesses rather than victims, it difficult to determine how personal relevance affects juror perception. The exception is Pozzulo and Dempsey (2009), who found that there was no effect of age on credibility when the individual was a victim, but children were perceived as less credible than adults when they were bystanders. Also, child victims were perceived as more credible than child bystanders.

**Limitations**

After analyzing the existing research on juror perception of child eyewitnesses and integrating it with what we know about childhood memory, several shortcomings and holes in the literature have become apparent. Some of these are: the psychology field’s emphasis on trends given the need for individual assessments in the legal system, the modality of the testimonies included in the analyzed experiments, the ages of the witnesses in the experiments, the demographics of the mock jury members, and the types of crimes included in the experiments.

One important limitation to consider when making recommendations is that psychology looks at trends, while the legal system focuses on individuals. Although it
is useful to understand the strengths and weaknesses of children’s memory at different stages of development, as well as the various factors that affect their testimony, it is important to remember that many of these factors are subjective. Children develop at different rates, and under different conditions. For this reason, it is extremely difficult, and even harmful, to make a blanket statement about the reliability of children at a given age. The recommendations included below are based on laboratory studies and should be interpreted as guidelines. It is important to always consider the specifics of the case (including the child’s personal attributes) when reaching a verdict.

Also related to the differences between the experiments included in this review and legal proceedings in an actual court is the format in which the testimony is presented. Many of the studies included here used written statements, even though such a testimony would generally be inadmissible in an actual courtroom. This is a limitation because it does not allow for the possibility that factors such as the child’s behavior and tone of voice will affect the jurors’ perception. The other modality commonly used in these studies is video recordings. These are more reflective of reality since children are sometimes allowed to testify via closed-circuit televisions. Nevertheless, the results of this analysis are still limited since none of the studies examined juror responses to in-person testimonies. Without understanding how jurors respond to in-person testimonies it is difficult to make recommendations that are applicable to real life situations.

A third limitation is the choice of ages that the various researchers used in their studies. To begin with, many of the studies compared children of specific ages
(e.g. four and nine) to “adults”. This catchall term makes it difficult to make comparisons between studies, and more importantly it makes it unclear at what age someone is considered an adult. Are researchers using the legal definition of an adult, meaning that anyone above the age of 18 is an adult? Or is it 22, the age when many people graduate university? Given that so many of the studies include age as a manipulation it is important to know who the children are being compared to. Also related to age is the fact that none of the studies look at teenagers between 14 and 17. In most of the studies, the cut-off for children is between 10 and 12, and then if they specify the age of the adults it usually starts up again between 18 and 21. As a result it is extremely unclear what jurors think of teenagers. Are children seen as being as reliable as adults after 12? What makes this gap in information even more troubling is the fact that teenagers and young adults experience a reminiscence bump between approximately 15-years-old and 30-years-old. During this period individuals retain a higher number of memories, and the memories contain a higher degree of detail. Despite the significance of this period during memory development, none of the studies I came across included teenagers. As a result we know little to nothing about their abilities to testify, and more importantly we do not know how jurors perceive them. Furthermore, there are a high number of individuals who pass through the juvenile court system each year, making it all the more important to understand how this demographic is perceived.

Along similar lines, another age related limitation is the demographic of mock jurors. Of the 23 experiments, 21 (91.3%) of the studies used mock juries made up entirely of university students. Of these 21, at least 17 of them used students either
primarily or exclusively from psychology classes. This is a major limitation because not only are university students in general not indicative of the wider population, but psychology students possess more knowledge about the factors involved in memory and eyewitness testimony than the average juror. As a result their judgment of child witnesses is likely to different. Two of the studies did include outside participants, but in one of these, over half of the participants were still university students (Krähenbühl (2012) mixed students and outsiders in the same experiments, while Goodman et al. (1987) conducted separate experiments).

A final limitation is the types of crimes involved in the experiments. None of the experiments involved recurrent crimes, or crimes that intersected with recurrent events in the children’s daily. As such it is difficult to determine the role that scripts play when children are recounting a crime. Without including these sorts of crimes it is also impossible to know whether jurors judge children differently depending on the frequency of the crime. Do they believe them more because they have witnessed the event multiple times? Or do the discount their testimony because they suspect it is an amalgamation of multiple events?

**Future directions**

One area where there is a clear need for further research is juror perception of teenagers. As mentioned above, in the limitations, the existing research on juror perception of child eyewitnesses focuses on children 12 and under. This may be because at a certain point jurors view teenagers as being as reliable as adults; however, without further research it is difficult to know when that point is. In
addition, there needs to be research on the effect of the “reminiscence bump” on eyewitness testimony.

Another way the research could be expanded is by studying the effect of in-person testimonies. As previously noted, video already has a much stronger impact than a written testimony or summery, and it would make sense that in-person testimony would have an even greater impact. To begin with, testifying in person would likely have an emotional impact on the child, which would in turn have an effect on the jurors. In addition, the act of being in the same room as the child, even if they are not emotional, might have an impact on the jurors’ judgments of the child.

As mentioned previously, both in the introduction and the discussion, maternal reminiscing style has an effect on autobiographical memory development. This form of memory is directly related to eyewitness testimony, and so it would be valuable to conduct a study that looks at how reminiscing style affects eyewitness testimony specifically. Also related to reminiscing, the current literature focuses on maternal reminiscing style, but it would be interesting to find out whether the same effect occurs if it is the father guiding the reminiscing. Logically, it should, but it is worth investigating. Furthermore, another avenue of investigation related to reminiscing is the impact on memory of reminiscing exercises in school. Is it possible for preschool teachers to guide young children in memory development exercises?

In addition to reminiscing style Fivush (2011) mentions that elaborative interviewing helps children recall more details as well as form more coherent memories. Therefore, it would be worthwhile to investigate the effect of interviewing style on eyewitness testimony, as well as the impact that testimonies produced
through different interviewing styles have on juror perception. There is already some research on questioning style (open vs. closed); but, as evidenced in the discussion, there is a lack of consensus at the moment about children’s abilities to answer these questions. As such the topic warrants further investigation.

Also related to interviewing is the issue of suggestibility. As previously mentioned, the power dynamic between children and authority figures can result in diminished reliability. An interesting question for future research is: are children more suggestible in terms of their actual memories, or do they simply change their memory reports in order to be compliant? In order to reduce the impact of this unbalanced relationship, we first have to have a clear understanding of its effects.

Finally, through this analysis we have seen that deliberation can have an important role in determining guilt as well as in judging the capabilities of the witness. As a result, it is important to delve deeper into the role of deliberation in decision-making, as well as investigate how the personalities of individual members of the jury affect deliberation. After all, it would be a problem if a single individual with a strong personality dictated the responses of an entire jury.

Conclusions

In writing this paper I set out with two aims. The first was to identify what research already exists on juror perception of child witnesses and analyze it in relation to what we know about child memory. I feel confident that I was able to accomplish this. The second aim was to develop a set of recommendations for jurors who are presented with a child witness. Given the current state of the field, I was only
partially able to achieve this second aim. The main takeaways based on what we know now are:

- How jurors expect they will react is very different than how they actually respond to a child’s testimony.
- Modality has a major role in perception, with video being the most effective among the modalities included in this analysis.
- If jurors’ expectations are defied in a positive way (e.g. child appears very confident) that may override their negative biases, and they may give more weight to the testimony than they would otherwise.
- Children are able to form and convey coherent memories from about 5-years-old, but their narrative abilities are continually developing and improving. In addition, while they may lack details, the memories that children do have are often accurate, especially if the event has personal relevance.
- Crime type sometimes plays a role in how jurors perceive eyewitnesses. For instance, if the witness to a drug-deal is a young adult.
- Speech style, consistency, competence tests, etc. affect judgments as would be expected; however, there is sometimes an interaction between these measures and age that affects their degree of influence.

Even with all this evidence, there is still a lot of work that needs to be done before we can conclusively say how jurors should interpret children’s testimony. Given the present state of the literature, my recommendations are as follows:
1. If possible, children’s testimonies should be presented in-person, or on video at the very least.

2. Jurors should consider children’s age and speech abilities before judging their credibility, as young children are capable of coherent narratives.

3. It is important to consider the quality (accuracy, consistency, honesty) of testimony above the quantity (number of details)

4. Confidence may be a better predictor of perceived credibility than age.

5. Although it is important to remember that children develop at different speeds, there is evidence that between 9-years-old and 11-years-old children become able to recall autobiographical events at an adult-like level.
References


## Supplemental Table 1

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Crime</th>
<th>Testimony</th>
<th>Factor</th>
<th>Witness Age</th>
<th>Dependent Measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goodson, 1967</td>
<td>Fatal car crash</td>
<td>Transcript</td>
<td>Age</td>
<td>6, 10, 30</td>
<td>Degree of Guilt</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Credibility</td>
<td>Increased with age</td>
</tr>
<tr>
<td></td>
<td>Fatal car crash</td>
<td>Video</td>
<td>Age</td>
<td>6, 10, 30</td>
<td>Degree of Guilt</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Credibility</td>
<td>6 = 10, 10 = 36, 6 = 30</td>
</tr>
<tr>
<td></td>
<td>Murder</td>
<td>Transcript</td>
<td>Age</td>
<td>6, 10, 20</td>
<td>Degree of Guilt</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Credibility</td>
<td>6 = 10, 10 = 36, 6 = 30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Doubt (D)</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Manipulation (S)</td>
<td>6 = 10, 10 = 36, 6 = 30</td>
</tr>
<tr>
<td></td>
<td>Newcombe, 2007</td>
<td>Non-violent story</td>
<td>Video</td>
<td>Conflicting witnesses</td>
<td>4, 9, Adult</td>
<td>Accuracy Detection</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Same age</td>
<td>-4 v 9</td>
<td>Sig. above chance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Adult v adult</td>
<td>Sig. above chance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Non-violent story</td>
<td>Video</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Other's age</td>
<td>-4 v 9</td>
<td>Could not detect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Adult v adult</td>
<td>Sig. above chance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Other-accurate</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Adult v adult</td>
<td>Sig. above chance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sig. above chance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Liss, 1995</td>
<td>Abduction</td>
<td>Transcript</td>
<td>Age</td>
<td>8, 12, 20</td>
<td>Believability</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-High, moderate, low</td>
<td>Low = Moderate = High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Confidence</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy</td>
<td>Low = Moderate &lt; High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Still frame = Video</td>
<td>Low = Moderate &lt; High</td>
</tr>
<tr>
<td></td>
<td>Abduction</td>
<td>Video/Bill-frame</td>
<td>Age</td>
<td>8, 12, Adult</td>
<td>Believability</td>
<td>No effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-4 v 9</td>
<td>No effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Adult v adult</td>
<td>No effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Confidence</td>
<td>No effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy</td>
<td>No effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Still frame = Video</td>
<td>No effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Still frame = Video</td>
<td>No effect</td>
</tr>
<tr>
<td></td>
<td>Kellermeyer, 2012</td>
<td>Minor assault</td>
<td>Transcript</td>
<td>Age</td>
<td>4, 6</td>
<td>Child Characteristics</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Interviewer Behavior</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Child Understanding</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Case Progression</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td>Pazzola, 2006</td>
<td>Car theft</td>
<td>Transcript</td>
<td>Age</td>
<td>10, Adult</td>
<td>Credibility</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Verdict = Guilty</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Verdict Confidence</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td>Pazzola, 2009</td>
<td>Violent drug deal</td>
<td>Transcript</td>
<td>Age</td>
<td>10, 40</td>
<td>Credibility</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Victim</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Witness</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td>Peterson, 1996</td>
<td>Robbery-Murder</td>
<td>Transcri</td>
<td>Age</td>
<td>7, 20</td>
<td>Truthfulness</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy</td>
<td>No effect of sex</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20 = 7</td>
<td>No effect of sex</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No comp text = Comp text</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bruer, 2014</td>
<td>Murder</td>
<td>Transcript</td>
<td>Age</td>
<td>4, 12, 20</td>
<td>Verdict = Guilty</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Degree of Guilt</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Verdict Confidence</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Integrity</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Reliability</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6 = 12 = 20</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6 = 12 = 20</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6 = 12 = 20</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6 = 12 = 20</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td>Nwosuva, 2015</td>
<td>Non-violent theft</td>
<td>Video/Summary</td>
<td>Age</td>
<td>7, 10, 23</td>
<td>Cognitive Competence</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy of recall</td>
<td>23 = 7, 23 = 10, 7 = 10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Trustworthiness</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significance</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Summary</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Taped</td>
<td>No effect of age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Testimony Modality</td>
<td>Summary = Taped</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cognitive Competence</td>
<td>Summary = Taped</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy</td>
<td>Taped = Summary</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Trustworthiness</td>
<td>No effect of modality</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Taped = Summary</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Warning</td>
<td>No warning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>About</td>
<td>No warning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Children</td>
<td>No warning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Eyewitnesses</td>
<td>No warning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Verdict = Guilty</td>
<td>No warning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significanc</td>
<td>No warning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Summary</td>
<td>No warning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Taped</td>
<td>No warning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Summary</td>
<td>No warning</td>
</tr>
<tr>
<td>Source</td>
<td>Type</td>
<td>Age</td>
<td>Notes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>8, 21, 74</td>
<td>Accuracy Suggestibility Honesty Weight given to testimony 6:8 = 74 &lt; 21 21:74, 8 &lt; 6 6:8 = 74 &lt; 21 6:8 = 74 &lt; 21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leippe, 1989</td>
<td>-</td>
<td>8, 21, 74</td>
<td>Accuracy (Expected) + Negation 9 Reckel No effect of age 5:6 &lt; 8:9 &lt; 12-13 &lt; 19-22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Robbery-Murder Narrative</td>
<td>6, 10, 30</td>
<td>Verdicts - Guilty No main effect of age Weak = Ambiguous &lt; Strong Int: 6 = 10 most influential for strong case</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>6, 10, 30</td>
<td>Strength of evidence No effect of age Inconsistent = Consistent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>6, 10, 30</td>
<td>Verdicts - Guilty No effect of age No effect of consistency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>6, 10, 30</td>
<td>Credibility No effect of age Inconsistent = Consistent Int: 6 more credible when consistent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Robbery-Murder Transcript</td>
<td>6, 10, 30</td>
<td>Deliberation = Expected, Unexpected Strength of evidence 30 &lt; 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>6, 10, 30</td>
<td>Verdicts - Guilty 30 &lt; 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>6, 10, 30</td>
<td>Credibility 30 &lt; 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>6, 10, 30</td>
<td>Verdicts - Guilty (D) 30 &lt; 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>6, 10, 30</td>
<td>Credibility (D) 30 &lt; 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goodman, 1989</td>
<td>Shot-Medical Video</td>
<td>3-6</td>
<td>Accuracy Suggestibility Consistency Truthfulness Identification: accuracy Credibility Positively correlated with age No effect of age Positively correlated with age No effect of age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Credibility 25 &lt; 8 Powerless = Powerful No Int.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Degree of Gult (D) 25 &lt; 8 Powerless = Powerful. Qualified by int. Int: 1 = Powerful, 2 = Powerful</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Credibility (D) 25 &lt; 8 Powerless = Powerful No Int.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wells, 1989</td>
<td>Abduction Video</td>
<td>8, 12, 18-22</td>
<td>Credibility No effect of age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Confidence No effect of age</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>