Wilfrid Laurier University

Scholars Commons @ Laurier

Theses and Dissertations (Comprehensive)

2020

An Evaluation of the Timeline as a Memory Retrieval Aid for Children's Recall of Specific Details from a Repeated Event

Huan Huan Zhang
Wilfrid Laurier University, zhan8620@mylaurier.ca

Follow this and additional works at: https://scholars.wlu.ca/etd

 Part of the Child Psychology Commons, and the Developmental Psychology Commons

Recommended Citation

Zhang, Huan Huan, "An Evaluation of the Timeline as a Memory Retrieval Aid for Children's Recall of Specific Details from a Repeated Event" (2020). *Theses and Dissertations (Comprehensive)*. 2236.
https://scholars.wlu.ca/etd/2236

This Dissertation is brought to you for free and open access by Scholars Commons @ Laurier. It has been accepted for inclusion in Theses and Dissertations (Comprehensive) by an authorized administrator of Scholars Commons @ Laurier. For more information, please contact scholarscommons@wlu.ca.

AN EVALUATION OF THE TIMELINE AS A MEMORY RETRIEVAL AID FOR CHILDREN'S RECALL OF SPECIFIC DETAILS FROM A REPEATED EVENT

by

Huan Huan Zhang

Master of Science, National Taiwan University, 2015

DISSERTATION

Submitted to the Department of Psychology

in partial fulfillment of the requirements for

Doctor of Philosophy in Developmental Psychology

Wilfrid Laurier University

© Huan Huan Zhang 2019

Abstract

Investigators and prosecutors are heavily dependent on children's testimony in abuse cases where physical evidence is often lacking, making children the sole source of information. Decades of research have shown that young children are indeed capable of accurately recalling events from the past. Findings from research on interview techniques suggest that interview aids such as dolls and human diagrams are often not helpful and pose risks of eliciting inaccurate reports unless they are used cautiously and non-suggestively at the end of the interview.

The timeline, which is a visual depiction of time, is another type of interview aid that is sometimes used to elicit information about time. However, no clear evidence about its risks and benefits to children's recall has been established. This dissertation sought to answer two questions about the timeline in three experiments. First, does the timeline help children recall specific details about a repeated event and its respective temporal characteristics? Second, how do adults perceive the timeline as an interview aid in children's recall of temporal details?

Using a repeated-event paradigm, the first two studies examined children's recall of a repeated event when an interview used the timeline and without the timeline. The two studies also examined the effect of the timeline regarding two different types of interview questions. Namely, in the first study, children answered *Wh*- questions regarding a particular instance of a repeated event. Children in the timeline condition were less accurate and sometimes more suggestible than those in the control condition. There was no clear evidence in support of the visual aid. In Study 2, children's spontaneous recall of target details in a repeated event was analyzed. Using the same repeated-event paradigm, 6- to 9-year-olds were asked to respond to open-ended recall prompts (i.e., "*Tell me everything...*") with or without the timeline as a guide. The timeline did not lead to a more specific recall of the target items in the event. Results from

Study 1 and 2 were in line with previous studies that demonstrated interview aids are generally not helpful to children's recall and may pose risks of suggestibility.

Study 3 examined adults' perceptions of children's verbal recall and recall using a timeline. Two groups of adults watched two halves of a child's interview about a summer camp. One group watched the child interviewed without a timeline first and then with a timeline; the other group watched the interview in the reverse order. Adults gave ratings on interview characteristics and overall credibility after each half of the interview. They also rated how they perceived the timeline (e.g., helpfulness of the timeline). When participants had only watched half of the interview, the perceived overall credibility did not differ between those who watched the verbal and the timeline halves. After watching the second half of the interview, the overall credibility changed based on how adults perceived the timeline. Specifically, when the verbal interview was seen first, adults' change in the perceived credibility after seeing the timeline interview was positively correlated with their ratings of the timeline; when the timeline interview was seen first, their change in overall credibility after seeing the verbal interview was negatively correlated with their ratings of the timeline. Overall, participants did not differentiate children's credibility and other interview characteristics solely based on whether a timeline was used.

The current dissertation investigated the role of the timeline in the recall of details from repeated events. Collectively, findings suggest that the timelines used in the current studies do not provide additional benefit to children's verbal recall of repeated events. The risks associated with the timeline such as suggestibility and erroneous credibility perceptions also suggest that the timeline should be used with caution in investigative settings. Future research should focus on how to build upon a good verbal interview before introducing any interview aids.

Acknowledgements

Four years ago, when I started my doctoral studies at Laurier, I could not have imagined that 4 years later I would be able to achieve so much. This dissertation is evidence of my academic achievement, as well as my personal growth throughout this incredible journey. I could not have reached this milestone without the help of many, whom I would like to thank.

First, I would like to express my heartfelt gratitude for my supervisor, Dr. Kim Roberts. I have admired your work since the first time I read your article on child memory and witness testimony when I was in my Masters. I had a clear goal then that I wanted to continue doing child research in the area of your expertise. Thank you for accepting me into your lab 4 years ago. I truly appreciate the freedom you gave me on choosing my topic of interest, as well as the guidance on every step of the way.

Second, I would like to thank my committee members Drs. Alexandra Gottardo and Eileen Wood for all the suggestions, edits, and feedback that you provided on the many versions of my drafts. From the proposal meeting, to the pre-defence meeting, and to the defence, you both gave an incredible level of detail in your feedback. Also, thank you Dr. Andrew Welsh for your feedback and questions from a Criminology perspective. They were refreshing and invoked thoughts in areas that I had not thought about before. Thank you, Dr. Lindsay Malloy for your time and cooperation with the logistics of the defence. Your work and expertise in forensic psychology is truly inspiring.

I would not be where I am today without the support of my dear colleague and friend - Vivian Qi. I am tearing up as I am writing this and thinking back on the many good memories that we have shared throughout our four years of grad school together. Although we had our share of arguments, but they made our friendship stronger. I feel incredibly lucky to have shared

the office with you and had discussions and debates, even, on just about anything. I appreciate that you always respect my beliefs and care about my feelings. We witnessed each other's significant milestones and I am very grateful to have had you on this journey. I wish you all the best with your studies at UBC and look forward to our continued and forever friendship.

I would like to thank the research assistants that have helped me with data collection and coding - Julia Caggianiello, Monica Crnogorac, Andrea D'Alessandro, Raha Moradhasel, Riana Rajaram, and Hannah Cowan. Also, data collection in Taiwan could not have been successful without the help of Dr. Yee-San Teoh and the research assistants that were involved. Thank you to the past and current members of the Child Memory Lab. It was fun to have been around an energetic and fun group of young aspiring researchers.

Finally, I am grateful to my family and friends who always supported my decision to continue to study and supported me throughout challenging times. I am thankful for my parents who always give me the freedom to pursue what I want and respect my decisions. The path to becoming where I am today was not easy but knowing that my family is always behind my back on every decision that I make gives me incredible strength. Thanks to my little sister Kori whom I love very much, for being a good kid and giving joy and laughter to the whole family.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

General Introduction

Investigators and prosecutors are heavily dependent on children's testimony in abuse cases where physical evidence is often lacking, making children the sole source of information (Lamb, Orbach, Hershkowitz, Esplin, & Horowitz, 2007). For decades, researchers have been examining children's ability to recall memories from personal events in an effort to develop ways to improve their recall. Children are indeed capable of providing accurate and reliable accounts of what happened under circumstances where a good interview protocol is utilized (Lamb et al., 2007). However, the difficult task of retrieving memories for specific instances of a repeated crime remains challenging to both child victims and investigators. To tackle this problem, adult interviewers sometimes use interview aids to encourage children to provide more information.

Interview aids are objects used by adult interviewers in conjunction with verbal questioning, in an effort to close the gap between what children remember and what they are willing to tell (Poole & Bruck, 2012). In particular, extant research has been dedicated to examining whether and how anatomical dolls and human figure diagrams may influence children's recall (e.g., Aldridge et al., 2004; Teoh, Yang, Lamb, & Larsson, 2010). Anatomical dolls are gender-differentiated dolls that depict all of the primary and secondary sexual characteristics of a human. Evidence supporting, as well as cautioning, the use of these interview aids varies across studies (see Poole & Bruck, 2012 for a review). Nonetheless, there is a wide consensus on limiting the use of visual aids in child interviews until after a thorough verbal interview is conducted (Lamb, Hershkowitz, Orbach, & Esplin, 2008). Research on the effect of interview aids on recall has also expanded to include other interview aids such as free drawing

(e.g., Otgaar, van Ansem, Pauw, & Horselenberg, 2016). Results suggest that making a drawing about what happened helped increase the completeness of children's recall but at the expense of an increase in false memories compared to a control group (e.g., Bruck, Ceci, & Francoeur, 2000; Otgaar et al., 2016). In recent years, the timeline, which is a graphical depiction of time, has entered the arena of research on interview aids (e.g., Gosse & Roberts, 2013). Police and social workers have expressed that they occasionally use a timeline in their questioning as it is believed to be a helpful addition to child interviews. However, evidence regarding the influence of timelines on children's recall is not yet known. Similar to the research paths of anatomical dolls and body diagrams, a crucial step to understanding the impact of the timeline on children's recall is for it to be evaluated before it can be confidently included in forensic interviews.

Roadmap

The current program of research examined the role of the timeline in children's recall, as well as adults' perceptions of children using this visual aid. The first chapter provides an overview of children's memory development for events. Sections that follow set the context for the research questions by: a) first describing the forensic relevance of children's memory research and the challenges children face as eyewitnesses, b) understanding the importance of studying repeated-event memory and children's difficulties in recalling specific instances of a recurring event, c) describing the origins of interview aids and their current state in empirical research as well as their usage in the field, d) providing an overview of children's temporal development, and e) discussing the potential benefits and risks of the timeline with respect to other interview aids and children's development. Throughout this introductory chapter, reference is made to events/experiences identified as the Laurier Activities. The Laurier Activities refer to a repeated-event paradigm used in previous repeated-event studies (e.g., Brubacher, Glisic,

Roberts, & Powell, 2011), as well as in Study 1 and Study 2 of this dissertation. In these repeated events, children participate in four group activities such as building a puzzle and reading a story, each lasting approximately 20 minutes. The structure of each day of the Laurier Activities is identical (e.g., always build a puzzle and read a story), however, details specific to each day vary across the four occurrences (e.g., a snowman puzzle on occurrence 1 versus a clown puzzle on occurrence 2). In addition, a unique badge/bracelet (e.g., a candy badge) is typically given to children in one target occurrence of the Laurier Activities and later used in interviews as a reference for the target occurrence (e.g., *"What was the puzzle you built the time you got a candy badge?"*). The details of the Laurier Activities and the design used in the present studies are described in full detail in Chapter 2.

Memory development

Evidence for memory can be seen hours after babies are born. Newborns are capable of recognizing their mothers' faces shortly after birth (Bushnell, 2001; Field, Cohen, Garcia, & Greenberg, 1984). Two- to 6-months-old infants are capable of learning through reinforcement by kicking their feet to activate a mobile in their crib (Rovee & Rovee, 1969). Before children learn how to talk, their memories are encoded in nonverbal modalities such as through visual or tactile means (Hayne & Rovee-Collier, 1995). Although these memories from infancy are fragile and difficult to retrieve later in life, evidence suggests that they are accessible at a young age through means of re-enactment or by using props, but not verbally (Simcock & Hayne, 2002). Once language skills develop, children can verbally demonstrate their memories for past events (Nelson, 1993). The rapid increase in language comprehension and production in the toddler years provides a new way for children to acquire and express their memories (Fenson et al., 1994; Hayne & Simcock, 2009). As such, at around 3- to 4-years-old, they begin to recall past

experiences in narrative form (Fivush & Nelson, 2004). Some researchers posit that the development of language skills is accompanied by the onset of a knowledge structure that organizes memories from previous experiences that happened to the self, and this is known as the cognitive self (Howe & Courage, 1993, 1997). Together with the onset of other cognitive functions, the development of the cognitive self leads to the emergence of autobiographical memory. Autobiographical memory contains information about the time and place of one's personal past and emerges around preschool years (Nelson & Fivush, 2004). When children begin to form autobiographical memories, they assume the role of narrators of their own past. For example, children become the source for sharing their experience in educational and social settings, and unfortunately, sometimes in legal settings. Their memories of past events in abuse cases are crucial because their testimony may be the most important piece of evidence. In an overview, the following section reviews the development of children's memory for events in general. Later, memory for repeated events, in particular, is discussed for its unique processes and characteristics, and relevance in the forensic context.

Episodic memory also serves as the foundation for autobiographical memory (Nelson & Fivush, 2004). Remembering details that describe the "*what"*, "*where*", and "*when*" about previously experienced events at a particular time and place form episodic memories (Clayton & Dickinson, 1998; Tulving, 2002). According to Tulving (2002), episodic memory is a memory system that allows humans to engage in mental time travel to the past when an event was experienced, and re-experience that previous experience, which subsequently allows for the retrieval of memories from specific episodes about the past. Memories for any particular event have been assessed using measures such as recall and recognition. Recall memory involves actively retrieving details of past events with or without cues, sometimes through narratives.

Recognition memory involves identifying whether or not the information provided belongs to previously experienced events (e.g., Ghetti & Lee, 2011). Children's narrative recall of events can be brief and skeletal (see Pipe, Thierry, & Lamb, 2007 for reviews; Powell, Fisher, & Wright, 2005), but it is evident that these narratives become increasingly informative with age (e.g., Brubacher et al., 2011). This age effect could be attributed to children developing more advanced language skills, knowledge and understanding of events and the self, and memory retrieval skills (Lamb et al., 2008). Kulkofsky, Wang, and Ceci (2008) found age was significantly correlated with the quality of 2- to 5-year-olds' narratives about a target event. The researchers posited that the quality of narratives likely reflects the strength of memory traces retained. Stronger memory traces reflect memories that are retained through a close-knit network of "richly associated" representations that include "semantic and formal features". Elaborated and cohesive narrative accounts can be recalled when those foundations are present (Pezdek & Roe, 1995). Recognition memory in the broader context has been assessed through various recognition tasks such as identifying studied words or phrases from a list (e.g., Craik & McDowd, 1987) or identifying pictures of faces that have been seen (e.g., Sophian & Stigler, 1981). In investigative contexts, recognition memory for events has been studied often through forced-choice and yes/no questions (e.g., Lamb et al., 2007). Sophian and Stigler (1981) posited that the development of recognition memory should be stable and that any age differences could be attributed to differences in the use of verbal labels, decision criteria measured through signal detection analyses, and/or skills in visual scanning.

Information gathered through recall and recognition memory assessments provides the key to the "*what*", "*where*", and "*when*" of an event, which serves as the pillar to any criminal investigation where the witness testimony provides the critical evidence. In eyewitness literature,

various types of recall and recognition questions have been used to assess children's memories of an event. Recall questions include invitational prompts (e.g., *"Tell me everything …"*), cued recall questions that use previously provided information to further elicit details (*"You said X, tell me more about X"*), or focused-recall questions that request information from a specific category (e.g., *"What was the color of …"*). Rather than using recognition questions, invitational prompts and recall questions are strongly encouraged in order to retrieve uncontaminated accounts (Lamb et al., 2007). On the other hand, recognition questions, or option-posing questions are not recommended unless crucial information is not addressed by recall prompts (Lamb et al., 2007). Option-posing questions require children to make identifications from a list of previously undisclosed information, which poses risks of misleading witnesses with false information. Brubacher et al. (2011) found that although open-ended recall narratives tended to be briefer for younger than older children, 4- to 12-year-old children's reports were equally accurate. Recall questions reliably elicit more informative reports compared to recognition questions (e.g., Sternberg, Lamb, Orbach, Esplin, & Mitchell, 2001). Furthermore, since recognition questions often consist of information not previously disclosed by children (e.g., "*Did the man have a red hat on*?" when the child had not said anything about "the man" or "hat"), previous literature has demonstrated that this may pose risks of suggestibility and false memories for children in general, and may be riskier for children under 6 (see Lamb et al., 2007 for a review).

Repeated-event memory

Research has demonstrated that recalling specific instances from a repeated event is an immensely difficult task, especially for children (Brubacher et al., 2011). However, in many jurisdictions, providing details specific to one instance of abuse when the abuse has happened

multiple times is necessary for a successful prosecution (Guadagno, Powell, & Wright, 2006). Also, testimonies that include more episodic details about specific instances than generic details are often perceived as more credible (Schneider, Price, Roberts, & Hedrick, 2011). Together, these requirements support the importance of identifying protocols that will increase the likelihood of children recalling more information, and more specific details when testifying about abuse. It is also crucial that investigators are equipped with knowledge regarding the development of memories and protocols for a repeated event in order to elicit episodic details that point to unique occurrences, a process known as particularization (Guadagno et al., 2006).

Memories from repeated events differ from those of one-time events in the way they are formed and retained. Particularization requires one to make decisions about the origins of those memories, known as the source, which can take on many forms such as location, medium, person, or time (Johnson, Hashtroudi, & Lindsay, 1993). The process of making source decisions is known as source monitoring (Johnson et al., 1993). People engage in repeated events such as going to work, buying groceries, and acquiring new information, and thus are constantly monitoring the source of those memories in their everyday lives. Particularly, in investigative settings, the temporal source of remembered information is one of the key references that an eyewitness report needs to include in order for the allegation to be substantiated (Lyon & Saywitz, 2006; McWilliams, Lyon, & Quas, 2016).

Repeated events are a regular part of everyday life. For example, parents and their children might typically go to a supermarket on the weekend to buy groceries. During each grocery shopping experience, they might get a shopping cart, go through the vegetable and fruits section, pass the meat and dairies, perhaps pick up a frozen pizza, contemplate whether to get potato chips and soda and finally arrive at the check-out. Memories for these repeated activities

are often discussed in terms of "scripts" (Schank and Abelson (1977) and are a structured set of sequences of events made up of "slots and requirements" which can be filled with specific details. The essence of scripts is that they are highly structured and contain details that are invariant. Referring back to the grocery shopping example, a script for a weekly shopping trip may consist of going to the supermarket, getting a cart, passing various sections, and arriving at check-out. While the structure remains the same each time, each "slot" can be filled with specific details. In addition, any given script is causally and temporally linked, where one sequence of events or activities enables the occurrence of the next sequence by a causal or temporal relationship. According to Schank and Abelson (1977), specific memories of a recurring experience can be established throughout a script and stored as "pointers"; then the script can become available by simply remembering these pointers rather than the entire script. As such, causal and temporal links may act as cues at recall (e.g., because the meat and dairy sections are in the back of the supermarket, one would remember that going through the fruits and vegetable section happened before getting a carton of milk).

Developmental psychologists later examined children's development of event memories for recurring events, making references to script development (e.g., Fivush, 1984; Hudson, Fivush, & Kuebli, 1992; Hudson & Nelson, 1986). According to Hudson et al. (1992), children's recall of a recurring event is characterized by, 1) a lack of memories for specific instances, 2) a generic reporting of a set of expectations that each instance will happen (e.g., "*You do puzzles at the camp*"), and 3) organizing details temporally according to the order of events that happened. Children as young as 3 years of age can recall general features about a recurring event such as "*You buy things and then you go home*" in a grocery shopping experience (Hudson et al., 1992; Nelson, Fivush, Hudson, & Lucariello, 1983). However, Brubacher, Roberts, and Powell (2012)

demonstrated that children recalled the script for a repeated experience with less episodic language than those who had a single experience. With regard to the temporal sequence of script recall, Farrar and Boyer-Pennington (1999) found 4-year-olds could sequence correctly only 8% of the actions in an event (Experiment 1) and increased dramatically to 61% in Experiment 2. This difference was largely due to the number of details remembered in each experiment. This finding also indicates that the ability to sequence details within a script is not a consistent trait of recalling a script. However, Hudson and Nelson's (1992) characterization of scripts may be based on real-life events where actions tend to be interconnected and where one detail affects another (Schank & Abelson, 1977), whereas Farrar and Boyer-Pennington (1999) used a novel event in a laboratory setting where activities may not be closely knit in terms of their interconnectedness (7 activities about magic). The studies in this dissertation utilized novel repeated-event paradigms that included activities and details that are not as meaningfully connected as those in the natural setting (e.g., build puzzles and make drawings). As such, it may limit children's ability to sequence details in the scripts developed for the repeated event memories.

Nelson (1988) posits that multiple episodes are incorporated into a script, separating details of specific instances from the script. While an individual can retain both a generic and episodic representation about event experiences (Hudson & Nelson, 1986), children's event schemas are dominated by representations of event experiences that predominately focus on generic scripts (Nelson, 1988). Despite having difficulties recalling specific episodes compared to generic scripts, researchers agree that children are capable of recalling episodically when the interviewer poses questions in an episodic manner (e.g., Brubacher, Malloy, Lamb, & Roberts,

2013; Brubacher, Peterson, La Rooy, Dickinson, & Poole, 2019; Brubacher et al., 2012; Schneider et al., 2011).

Goodman's (1980) original schema-deployment confirmation model suggests that the script-formation process is based on confirming details that belong to the script, Farrar and Goodman (1990) further posit that this process differs across developmental phases depending on the level of schema development. In the schema-confirmation phase where a schema for an event is still in development, incoming information will likely be incorporated into this developing representation. For example, details from a single party-going experience will likely be fully incorporated into the party-going schema; no detail from that experience will be separated into episodic memory. Once a mental representation of an event has developed into a coherent schema, incoming information that does not match such representations will likely be attended to on a selective basis. This schema-deployment process triggers the development of a second representation in episodic memory rather than in generic memory where the script exists. For example, after having built a schema for attending parties, a unique experience that is different from the usual, such as having a drink accidentally poured on you, will likely be placed into episodic memory.

Children develop more elaborate scripted knowledge as they gain more exposure to recurring experiences such as going to birthday parties and participating in snack times every day at school (Hudson et al., 1992). According to the schema-deployment confirmation model, younger children may require more experience and a longer process than older children and adults in forming a schema, and thus have difficulties separating discrepant event information (episodic details) from a general representation (script). At 2 ½ years-old, children tend to recall details that are routine even though details that are novel and do not fit into the typical script

occurrence are more salient and exciting than routine (Fivush & Hamond, 1990). As younger children depend more on the general event representation and retain information that is consistent with that representation, they are more likely than older children to have difficulty remembering specific event episodes (Farrar & Goodman, 1992). Thus, children increasingly tend to recall more specific and atypical information that deviates from the script as they grow older (Hamond & Fivush, 1991). In addition, older children recall a more elaborated and complex script that includes generic information using conditional and optional actions (Slackman, Hudson, & Fivush, 1986). For example, script recall could include conditions under which a script sequence is triggered (e.g., *"If A happens, then we do B"*) or what *could* happen (e.g., *"We could do puzzles at Laurier Activities"*). Moreover, younger children's rudimentary and less elaborated verbal reports are not solely due to their having less repeated exposure to knowledge and experience (e.g., Farrar & Goodman, 1992). Various repeated-event memory studies have shown that the same amount of experience led to younger children reporting fewer generic details than their older peers (e.g., Brubacher et al., 2012). Moreover, children's recall of a recurring event tends to include general script-like responses such as *"We always build a puzzle at the Laurier Activities"* (e.g., Brubacher et al., 2011; Brubacher et al., 2012).

Fuzzy trace theory posits that memories consist of verbatim and gist traces that are encoded simultaneously (Brainerd & Reyna, 1990). For repeated-event memories, verbatim traces are specific details about an event that include episodic details (e.g., *"We built a clown puzzle when we wore a candy badge"*). On the other hand, gist traces refer to a general gist of what happened (e.g., *"We always build a puzzle at the Laurier Activities"*). Memories for verbatim traces decay much more rapidly for younger children while gist traces strengthen after each experience (Brainerd & Reyna, 1990, 2004). As findings from studies using a repeated-

event paradigm have shown, memories for details that are always present are often highly accurate (e.g., Brubacher et al., 2012; Powell & Thomson, 1996), as each occurrence of the detail acts as a rehearsal, which strengthens its memory traces. On the other hand, details that change across multiple occurrences are often misattributed to the incorrect occurrence (i.e., weakened verbatim traces), which is a type of source-monitoring error commonly referred to as internal intrusion errors (Brubacher et al., 2012; Powell, Roberts, Ceci, & Hembrooke, 1999; Powell & Thomson, 1996). In a forensic context, when reports contain details that are inconsistent and misattributed to the wrong occurrence, the perceived credibility of the child may be undermined (Connolly, Price, Lavoie, & Gordon, 2008).

Children's memory of repeated events has been studied in laboratory settings. These repeated-event memories have been compared to single event memories (Powell & Roberts, 2002; Woiwod, Fitzgerald, Sheahan, Price, & Connolly, 2019), as well as their unique characteristics (Brubacher et al., 2011). Connolly and Lindsay (2001) found children to be highly suggestible when responding to recognition (yes/no) questions after repeated experiences compared to a single experience. A similar suggestibility effect was seen in Powell and Roberts' (2002) study that showed children with repeated experiences were more suggestible and gave fewer correct answers than those in a single-experience condition. Several types of details in a repeated experience help explain why memories from repeated experience are poorly recalled. *Fixed* details are those that remain consistent throughout a repeated event. In the grocery shopping example, fixed details may be the mode of transportation for each trip where one would always drive to the supermarket. Fixed details in repeated experiences tend to be remembered better than details that vary by 3- to 9-year-olds because they appear in consistent manners for each occurrence (Brubacher et al., 2011; Powell et al., 1999). *Variable* details are those that

occur regularly in the repeated experience but in different versions such that they vary across occurrences of a repeated event. Instead of driving to the supermarket, the mode of transportation varies each week where one could take the bus in week 1, taxi in week 2, walk in week 3, and so on. These versions of the detail are referred to as instantiations (Powell et al., 1996). Other non-fixed details in a repeated experience include high (*hi*) and low (*lo*) frequency details, and *new* details that only appear once. Refer to Table 1 for the types of details included in the current program of studies. The difficulty in attributing instantiations to the correct occurrence for children under 9-years-old has been well documented. Brubacher et al. (2011) found that 4- to 8-year-olds identified the correct occurrence for *hi* items with higher accuracy than *new*, *variable*, and *lo* items; and *new* items were better remembered than *lo* items. In addition, when details that varied in a repeated experience (*variable* details) have high similarity, they trigger unique verbatim traces rather than the same gist, this leads to a higher chance for false reports (e.g., Connolly & Price, 2006; Powell, Roberts, & Thomson, 2000).

Age effects for repeated-event memory are posited by both Fuzzy-Trace and script theories. Both theories suggest that memory for specific instances is weakened as time goes on, more so in younger children than older children (Woiwod et al., 2019). Specifically, Fuzzy-Trace theory notes that younger children's memories for verbatim traces deteriorate more rapidly than older children (Brainerd & Reyna, 1998; Brainerd, Reyna, & Forrest, 2002). Also, as scripts are formed as events repeat, script theory notes that older children are more sensitive than younger children to changes that occurred in the event script (Woiwod et al., 2019). Studies consistently find that older children outperform their younger peers with respect to repeated-event recall accuracy. For example, 4-year-olds provided more incorrect recognition responses to suggestive questions regarding a repeated event than 8-year-olds (Connolly & Lindsay, 2001). Brubacher et

al. (2011) found 7- to 8-year-olds to be better than 4- to 5-year-olds in making the correct attribution for non-fixed details.

Children as witnesses

Children's involvement in legal settings often relates to abuse and leads to them becoming witnesses. The World Health Organization describes the dynamics of child sexual abuse as one that, "… often occurs as repeated episodes that become more invasive with time" and may happen over weeks or years (World Health Organization, 2003). A recent Ontario Incidence Study Report revealed a stunning 68% of more than 848 child sexual abuse cases involve multiple incidents (Fallon et al., 2015). In Canada, where the legal system stemmed from British Common Law, children are required to provide details that make references to time, place, and other contextual information regarding specific occurrences so that specific charges can be laid (Guadagno et al., 2006). Errors or failure to provide this information gives a fair opportunity for the defendant to counter those charges. As such, research on memories of a repeated event is essential in understanding a fuller picture of children's recall abilities in a forensic setting. Due to developmental limitations, children and investigators face immense difficulties bringing forensically meaningful testimonies forward in the legal process.

Interview protocols. Researchers and forensic experts recognize the difficulties child witnesses and adult interviewers face in investigative interviews. Concerns about inappropriate and suggestive interview techniques used in high-profile child sexual abuse cases in the 1980s prompted researchers to develop interview protocols that best suit children's social and cognitive developmental needs. For example, the RATAC® Protocol developed by Children's Advocacy Centre in 1989 encourages spontaneity in children's responses, rapport-building, anatomy identification, touch inquiry, and abuse scenario (Anderson et al., 2010). While this protocol has

been adopted by many jurisdictions in the United States, no research has been published regarding its implementation and effects (Lamb et al., 2007). On the other hand, a protocol that has been widely studied by forensic experts is the protocol developed by The National Institute of Child Health and Human Development (NICHD). The NICHD protocol was developed to help both child witnesses and investigative interviewers (Lamb et al., 2007; Orbach et al., 2000). The NICHD protocol is an evidence-based investigative interview protocol that takes into account children's cognitive and social developmental capabilities and aims to increase the forensic relevance and accuracy of children's accounts while minimizing risks of false reporting (La Rooy et al., 2015).

Under the guidelines of the NICHD protocol, interviewers are encouraged to start the substantive phase of the interview with a general invitational prompt (*"Tell me what happened ...")* and follow up with information provided by the child (Lamb et al., 2007). Eliciting recall memory using open prompts does not constrain children's memory to any specific topic, but rather allows children to retrieve memories that are most accessible (La Rooy et al., 2015). Hence, memories reported in response to open prompts tend to be longer, more detailed, and more accurate (e.g., Lamb, Hershkowitz, Sternberg, Boat, & Everson, 1996; Sternberg et al., 1996). While using open prompts could elicit the most accurate memories, details from these memories are not always enough for investigative purposes (i.e., do not include critical forensic-relevant information for a case to proceed). Following exhaustive recall through open prompts, focused-recall questions that address specific details about previously mentioned details are recommended. These focused-recall questions such as *"When did [mentioned event] happen?"*, request information from specific categories (e.g., time). The NICHD protocol emphasizes using previously mentioned details, rather than non-mentioned

details that could be suggestive. While asking these focused-recall questions addresses the issue of missing critical information in the free-recall, the accuracy of the focused-recall questions tend to be lower than that of free-recall (Orbach et al., 2000).

The benefits of the NICHD protocol have been well-documented through many field studies including enhanced eyewitness reports from child witnesses and improved interviewer behaviors (e.g., Cyr & Lamb, 2009; Orbach et al., 2000; Sternberg et al., 2001). The protocol is now widely accepted and used in many countries and has been translated into nine languages (La Rooy et al., 2015). Despite the many evidence-based recommendations and procedures on how to interview young children, the NICHD protocol does not explicitly recommend the use of visual aids due to a vast number of studies that show mixed and conditional findings regarding their use. Most of these studies examined typically-used interview aids in child sexual abuse cases, namely anatomical dolls and body diagrams, as they seemed fitting to the discussion of body touch. While many of these findings regarding the use of dolls and body diagrams share common recommendations with that of other visual aids such as drawing (e.g., Otgaar et al., 2016), efforts into examining other interview aids should be continued to ensure that the benefits and risks unique to those props are not overlooked. The timeline may be considered most fitting in recalling repeated events which includes multiple time points and episodes. The following section reviews the current state of literature in interview aids and describes the current recommendations and their associated risks.

Interview aids

A recent field study conducted in New Zealand revealed that 62% of 98 investigative interviews with children involved the use of at least one visual aid (Wolfman, Brown, & Jose, 2018). Although statistics suggest visual aids are still popular among forensic interviewers,

findings from decades of research examining their efficacy on children's report accuracy vary across different types of aids with most researchers suggesting that they are not particularly helpful in the investigative context (e.g., Salmon, Pipe, Malloy, & Mackay, 2012).

In the 1980s and 1990s, visual aids were widely adopted by child-protection professionals as an important advance in communication techniques when interviewing child victims (Whitcomb, 1991). Despite their popularity, there are mixed results surrounding the use of these non-verbal props, as well as results showing limited but conditional support for their usage in investigative settings. Findings from a field study by Teoh et al. (2010) suggest that 4- to 13-year-old children who alleged sexual abuse were able to clarify and elaborate previously reported details using human figure diagrams but only after an exhaustive verbal interview. In another field study, Katz and Hershkowitz (2009) examined the effects of drawing "what happened to you" in interviews with alleged victims of sexual abuse aged between 4 and 14 years old. The children were first interviewed verbally using the NICHD protocol, and then randomly assigned to one of two groups: with drawing and no drawing. In both groups, children were asked to tell the interviewer again "*everything that happened to [them] from the beginning to the end as best as you can*". Children in the drawing group did so after they had completed their drawings. The researchers found that children in the drawing group provided more information via recall about the abusive event that included central details of the incidents. Results from this study suggested that free drawing may be an effective aid to enhance the amount of information children recall in abuse investigations (Katz & Hershkowitz, 2009).

In a laboratory study, Higgs (2013) explored the effects of using children's own photographs as an aid when interviewing children about a touch experience. The rationale was that, as using human figure diagrams requires cognitive competence in symbolic representation

(DeLoache & Marzolf, 1995), by using a picture of the child him/herself, such cognitive load may be reduced, and at the same time serve as a more effective retrieval cue (Higgs, 2013). Findings from DeLoache (2000) also suggest ways that could reduce this demand by showing that when 2 ½ to 3-year-olds are distanced from physically accessing a representational object (i.e., a room model placed inside a glass box, pictures, or videos), the children were able to better use the objects as symbols and succeed in the standard-model task (DeLoache, 1987, 1991). Using a human figure drawing or photograph after exhaustive verbal interviews produced more information about a touch experience compared to not using a prop; No difference was found between the human figure drawing and photograph interviews. A more recent study by (Bruck, Kelley, & Poole, 2016) suggests that using focused-recall questions (*Wh*- questions) and recognition questions (yes/no questions) in combination with body diagrams increased the number of correct responses regarding touch experience. Otgaar et al. (2016) also found evidence to suggest that when 6- to 12-year-olds were asked to draw and tell about what they could remember about an event, the completeness of their reports of the event was increased. However, it should be noted that both the Otgaar et al. (2016) and Bruck et al. (2016) studies found positive effects of visual aids at the expense of accuracy. While these positive, but limited, findings from both field and laboratory research suggest visual aids could be helpful to children's recall in some way, the findings also indicate they should be used with great caution.

The positive impacts of visual aids in some research are not universal. Specifically, several studies found that human figure diagrams do not facilitate children's reports of touch (e.g., Poole & Dickinson, 2011; Salmon et al., 2012; Willcock, Morgan, & Hayne, 2006). Further, Poole and Dickinson (2011) found that the timing of the introduction of body diagrams affects children's reports. Four- to 9-year-old children participated in science demonstrations where half

of them were touched innocuously. Their parents later read stories to them with accurate and inaccurate information about the demonstration. Results indicated that body diagrams led children to be more suggestible (i.e., reporting inaccurate information about their experience) than verbal questions alone, and proposed that the practice of introducing drawings early in interviews should be restricted. In addition, the authors suggest that by introducing the names of certain body parts in the pre-substantive phase using a body diagram (prior to asking about target event) led more children to misreport touches they had not experienced during the event. In another study, Brown, Pipe, Lewis, Lamb, and Orbach (2007) found that children who were interviewed with a human figure drawing produced a greater number of incorrect details than those only interviewed verbally. The number of forensically relevant errors did not vary by interview condition. These findings collectively suggest that non-verbal props should always be used with caution as they can be suggestive in nature and are often associated with more risks than any potential benefits they may bring.

In addition to using human figure drawings and photographs, anatomical dolls gained popularity amongst investigative interviewers in the 1970s as they were seen as a tool to help evaluate and extract critical information from young children involved in alleged sexual abuse cases (Poole & Bruck, 2012). They quickly gained popularity as can be seen in prominent sexual abuse cases in the United States (e.g., McMartin Preschool Trial; State vs. Kelly). In the case of the McMartin Preschool Trial, criticisms arose surrounding the use of anatomically correct dolls in combination with “leading questions and subtle pressures” (Linder, 2007). In addition, the children who were interviewed with dolls produced bizarre, inconsistent, and contradicting testimony in the preliminary hearing. This and other alleged sexual abuse cases soon sparked a

series of field and laboratory studies that examined whether anatomical dolls are beneficial or harmful to children's eyewitness accounts.

As many researchers suggested, dolls may serve as a communication tool for younger children as they may support their still emerging language skills (Thierry, Lamb, Orbach, & Pipe, 2005). However, some researchers have observed play behaviors unrelated to a target event when children were presented with the doll. Bruck, Francouer, Ceci, and Renick (1995) examined 3-year-old children's interactions with anatomical dolls by asking them to report their experience in a medical exam procedure. In their study, some children falsely demonstrated their experience of the medical exam through sexual actions with dolls such as placing their fingers inside the dolls' genitals. In addition to playful behaviors, researchers have also found that the use of dolls may lead to reports of imaginary details. In a field study, Thierry et al. (2005) found that, while the use of anatomical dolls elicited more fantastic details and ambiguous enactments than when they were not used, the amount and quality of information reported were not enhanced by the use of dolls especially in younger children (3- to 7-year-olds).

Understanding that objects have a representational meaning is essential to interview props being used appropriately in recall settings (e.g., body parts on a doll represent one's own body; maps represent locations; graphical timelines represent time). DeLoache and Marzolf (1995) discussed the issue of symbolic representation among young children and suggested that young children must first understand the dolls as a representation of themselves before proceeding to use dolls in forensic settings. In their study, 2- to 4-year-old children experienced activities that involved innocuous body contacts with a researcher. Findings indicated that although there was evidence that children understood that they were to use a doll to demonstrate what happened to them, they had difficulty mapping their body onto the doll. As a consequence,

their accuracy with dolls was lower than when they responded verbally or pointed to their own body parts. Bruck et al. (2000) also found that children's recall often included errors of commission when anatomical dolls were used (i.e., erroneous reports of touch that did not occur).

Timelines depicting a length of time as a horizontal line with visual markers of key times are often used in investigative settings to help children identify when events occurred. Field workers continue to use visual aids based on the long-held assumption that concrete objects may help children describe experiences that they otherwise would not be willing or capable of discussing (Poole & Bruck, 2012). Although commonly used, there is a lack of experimental or field studies about whether timelines truly enhance the quality of children's eyewitness reports. Therefore, two studies in this dissertation systematically examined the timeline as an interview aid.

Temporal memory development

As child sexual abuse cases are extremely complex in nature, developmental and cognitive barriers related to children's memory capacities create difficulties for investigators. Time is one of the crucial components of episodic memory (Tulving, 2002). Time information about an event helps identify the temporal nature of the event. Particularly, chronic and repeated abuse involves events that happen over a long stretch of time; over many days, months, or even years. As such, particularized reports that include when details occurred add substantial credibility to one's account (Schneider et al., 2011). However, young children often have difficulties providing such information due to their developing knowledge of temporal terms, linguistics abilities, as well as a general sense of time (e.g., Bauer & Mandler, 1992; Friedman, 1992). This section describes the development of children's temporal understanding and developmental trajectory with respect to the two dominant temporal memory development

frameworks, 1) memories from an event can be derived based on its distance in time from the present (distance-based theories), and 2) the locations of memories in established time patterns (location-based theories) (Friedman, 1993).

Research shows that children as young as two years of age begin to produce temporal terms such as “today” and “tomorrow” (Ames, 1946; Pawlak, Oehlrich, & Weist, 2006). However, this early onset of temporal language might not reflect children’s actual understanding of temporal meanings (Nelson, 1991). Harner (1975) demonstrated that 3-year-olds understood “today”; and that “yesterday” and “tomorrow” referred to non-present times, but did not understand that they refer to a specific time in the past or future. A recent study also demonstrated that 3- to 5-year-olds gradually develop the understanding for past, present, and future references, but were more accurate distinguishing between present and past actions than with present and future actions (Zhang & Hudson, 2018). In addition, young children show a partial knowledge of past and future understandings. For example, Tillman, Marghetis, Barner, and Srinivasan (2017) described a situation where one 21-month-old child Franny demonstrated an understanding of the term “tomorrow” by reacting correctly to her mother telling her about emptying the dishwasher tomorrow. As a 21-month-old, Franny also misused “yesterday” to represent events that happened at any time in the past, which is common among young children (e.g., Friedman & Kemp, 1998). As children acquire more temporal language, their temporal understanding for more complex and larger-scale time patterns such as days, years, and seasons develop gradually until adolescence (Friedman, 1986).

Distance-based processes lead to impressions of distance information about an event independent of knowledge about conventional time patterns (Friedman & Kemp, 1998). Through a series of studies, Friedman and colleagues showed that children as young as pre-school were

very accurate in judging the relative recency of two events (Friedman, 1991; Friedman, Gardner, & Zubin, 1995). Without the need for conventional time knowledge and a complex social and cognitive understanding of their surroundings, Friedman and Kemp (1998) suggest that children may be able to use such distance-based processes to judge the times of past events. The relativity-based theories explain the storing of new information in relation to older information stored in memory. Lewandowsky and Murdock Jr's (1989) theory of distributed associative memory, based on chaining models (See Brown, 1997 for a review), explained that the order of items is encoded based on the pairwise association between successive items. Each new item is added to memory and forms a chain made up of links between previous items in the list. However, Friedman (1993) pointed out that this process of numerous repetitive comparisons between items is rarely the case for episodic memory encoding. In a similar explanation, Hintzman, Summers, and Block's (1975) model posits that if new information enters memory storage and leads to the retrieval of an older memory trace, the order of the pair is automatically stored.

Location-based theories explain that memory for time is linked to locations in different scales of time (e.g., last week, the year of 2018) (Friedman, 1990). Some of the most adopted location-based theories are reconstructive theories, which posit that the process of arriving at the conclusion about time is drawn on knowledge about social, natural, and personal time patterns (Friedman, 1993). This process uses reconstructive methods to infer when events happened by combining aspects of memory with separate knowledge of time patterns (Friedman, 1993). According to Friedman and Lyon (2005), three components are necessary for the reconstruction of temporal memories: 1) the retrieval of temporal, episodic details; 2) a general knowledge of time patterns (e.g., days of the week, months); 3) and the cognitive ability to integrate episodic

memory and time patterns. For example, if details remembered about an event include wearing a full-sized winter jacket, you could derive from this clue that the event happened around wintertime. Friedman, Reese, and Dai (2011) showed that children's conventional time knowledge was related to the accuracy of their report of temporal information of events of time scales of up to 4 years. This finding suggests that possessing knowledge about general time patterns is a necessary step to the retrieval and accurate reporting of temporal information.

Clear developmental differences in children's ability to provide temporal information have been supported by experimental research. Findings support that children gradually develop knowledge for time patterns of different scales. Friedman (1986, Experiment 1) found that fourth-graders (Mean age = 9.9 years) could recite days of the week with high accuracy; and by fifth grade (Mean age = 10.7 years), children were able to judge months of the year (Experiment 2). Friedman and Lyon (2005) found that 4-year-olds were capable of reconstructing time for events that happened 7 weeks ago. Seven-year-olds performed poorly, however, when recalling events that happened 3 months ago, while 11- to 13-year-olds were within 40 minutes of the correct time of those events. While retention interval plays a role in determining children's ability to recall temporal information, developmental differences are also evident in these findings. In another study, Bauer, Burch, Scholin, and Güler (2007) asked 7- to 10-year-olds to self-nominate autobiographical memories. Children were highly accurate in judging their age and the season of when those events originally happened. This indicates that children under the age of 10 are capable of providing key temporal information about personally relevant events.

Furthermore, investigators often ask children to judge when events occurred in relation to landmarks such as birthdays or holidays as another way of obtaining time-relevant information from children (e.g., U.S. v. Tsinhnahijinnie, 1997). Findings from studies by Friedman and

colleagues showed that children under 9 years of age are capable of correctly judging the recency of two events with a distance of up to one year (Friedman et al., 1995; Friedman & Kemp, 1998). However, children often demonstrated a prospective bias, which is the tendency to proximate a forthcoming event as more recent than it actually is (McWilliams et al., 2016). Additionally, McWilliams et al. (2016) demonstrated that using recurring landmarks as references became more confusing for 6- to 10-year-old maltreated children. These children demonstrated prospective bias by making references to a forthcoming landmark as more recent rather than a past one. While the developmental trajectory of providing temporal memories is clear, the age at which children become fully competent in making temporal judgments on a variety of time scales is unclear.

The timeline as interview aid

Research for the timeline as an interview aid has surfaced in recent years. The timeline is a 2-dimensional graphical representation of time. Timelines are often used in educational settings as an instructional strategy to help children build a visual understanding of the past and develop concepts about time and continuity (Hoodless, 2002). For example, history lessons sometimes involve teaching historical events by placing those events on a visual timeline.

The timeline has been studied in the context of memory retrieval more often with adult samples. Hope, Mullis, and Gabbert (2013) asked adult participants to generate details related to persons and actions about the to-be-remembered event and to place those details on a timeline in order of their occurrence. Their findings suggest that the timeline facilitated adults' recall of the event by increasing the number of correct details and lowering sequencing errors at no cost to accuracy. In addition, Van Der Vaart and Glasner (2007) used a table that included time dimensions (e.g., years, months) and themes of data collected (e.g., job, domestic situation) to act

as a timeline to help adults recall autobiographical events. They found the timeline enhanced adults' recall of a past purchase experience and was especially helpful for a difficult recall task that involved less salient and less recent purchases. However, this table-style timeline is not the same as the graphical timelines used in the current study that displayed time anchors on a line.

Research on the effect of the timeline as a memory retrieval tool on children's memory of past events is scarce. However, previous literature that examined the development of temporal understanding has indirectly involved a timeline or a graphical representation of events in a temporal manner (e.g., ordering of events). For example, Friedman (1991) asked 5-, 7-, and 9-year-old children to judge the occurrence of two target events on different time scales. Each time scale was presented using a set of pictorial cards in order of their occurrence. For example, day-scale cards were represented by pictures of waking, eating lunch, eating dinner, and going to bed, seasonal scale cards were represented by pictures of fall, winter, spring, and summer. None of the children in that study were able to judge season-scale above chance; 7- and 9-year-olds were able to correctly judge using month-cards above chance; the 5- and 7-year-olds were above chance using day-cards. This study indicated that children gradually acquire temporal understanding on smaller time scales (i.e., day-scale) before moving onto larger time scales (i.e., season-scale). Although this study did not intend to test the effect of a timeline, in a way, the presentation of time cards served as a pictorial timeline. The findings suggest that young children may indeed benefit from a pictorial representation of time in aiding their memory retrieval of a past event, at least on a time scale representing shorter intervals such as days.

Using a timeline, Grant and Suddendorf (2009) examined 3- to 5-year-old children's ability to differentiate past and future events. The timeline was not the focus of the study, it only served as a testing tool. Children were asked to place target events that happened daily, annually,

and from several years ago or in the future on a timeline. Findings suggest that the 3- and 4-year-olds were able to accurately place recent past events and events that occurred during infancy on the corresponding location on the timeline. The 5-year-olds were able to distinguish both past and future events and make appropriate distinctions between recent and distant events. Although this was not a study on the effect of the timeline on the recall of past events, it provides some evidence that by age 5, children are capable of using the timeline to make temporal distinctions about the past and future.

Malloy (2002) compared 4- to 7-year-olds' verbal reports of temporal information and responses using a timeline. In response to a series of questions regarding the time of past events, children in the verbal condition were asked to report only verbally, and children in the timeline condition were asked to mark the event using a timeline. The accuracy of children's responses was determined by comparing children's responses to information provided by their parents. Overall, the accuracy in the verbal condition was not different from that of the timeline condition. This suggests that the timeline, similar to other visual aids, may not be any more beneficial than using verbal questions alone.

Research by Gosse and Roberts (2013) was one of the first published studies that directly examined children's ability to use the timeline to recall past events. In Experiment 1, parents of 4- to 8-year-old participants were asked to identify four unique and memorable events that happened in the previous week at distinctive times of the day that involved their children. These events were then marked on a horizontal timeline that corresponds to the time of day at which events occurred. Children's interviews began with an instruction about the timeline (e.g., the section between two anchors on the timeline represents an entire day). They were then instructed to place each of their parents' nominated events on the timeline that corresponds to when it

occurred. Findings from this study suggest that age played a role in children's ability to identify events using the visual aid. Specifically, the analyses of the recall of temporal location, order, and duration showed that older children's matching of events on the timeline was more consistent with their parents' estimates than 4-year-olds. These age effects are consistent with previous literature that demonstrated temporal understanding is not well developed until 7 to 8 years of age (Friedman & Lyon, 2005). Further, these findings were also in line with that of anatomical doll research suggesting that props may not work well with younger children (Poole, Bruck, & Pipe, 2011).

In Experiment 2, a different sample of children was asked to make general judgments to events nominated by the parents in Experiment 1. The purpose of this experiment was to examine whether children in Experiment 1 were genuinely using timelines to retrieve memories of the nominated events, or simply guessing and inferring when events occurred based on general knowledge. Results showed that children's judgments in Experiment 1 were closer to parents' estimates than those in Experiment 2, indicating that the children in Experiment 1 made genuine use of the timelines to help them retrieve memories. This study added to our understanding of children's use of a timeline to aid the retrieval of events that happened within a day (i.e., day scale). However, in child sexual abuse cases, especially repeated offenses, events span across days, weeks, even years. Therefore, it is important that we extend our knowledge of whether timelines aid children in identifying temporal information for events that stretch over a longer period.

The present program of research

The studies in this dissertation explored the role of the timeline in the investigative interviewing context using a repeated-event paradigm. Research in the last few decades has

mostly focused on two commonly used props in relation to child sexual abuse investigations; anatomical dolls and body diagrams. While there is research that finds evidence in support of the use of dolls and diagrams under some limited circumstances, there is a wider consensus that children's verbal recall should be exhausted before introducing props (Lamb et al., 2008). The effect of these aids has been discussed in relation to children's understanding of what these objects represent, as well as their abilities to allocate cognitive resources to the interactions. While little research has systematically examined children's use of the timeline, arguments and rationales for what is expected could be made analogous to that of dolls and body diagram studies. The first two studies in this dissertation utilized the repeated-event paradigm that contains four occurrences of the same set of activities and has been well demonstrated to reliably build a general event script (e.g., Brubacher et al., 2011).

The first study is an examination of whether the timeline is useful for 6- to 9-year-olds' recall of a specific instance of the repeated event compared to verbal recall alone. Specifically, children's ability to recall details in response to focused-recall questions (*Wh*- questions) about 16 target items from one specific instance of a repeated event was tested with and without the timeline. Refer to Table 2 and Table 1 for the target items and instantiations. In addition, children's ability to attribute details to the correct occurrence with and without the timeline was examined. Children in Study 1 participated in the Laurier Activities on four different days in a span of two weeks. Each occasion of the Laurier Activities included the same set of procedures (e.g., build a puzzle, play connect-the dots) but the details of those activities differed each day. This design allows for the analyses of the generic representation of the repeated event (what always happens) and episodic memories (what happened [the target instance] time). Children's accuracy was later assessed using focused-recall questions (e.g., *"What was the puzzle on [the*

target instance] day?"). Findings from Study 1 shed light on how the timeline compares to using no interview aid for children's recall accuracy.

While using interview aids can seem straightforward and appealing to child interviewers (Poole & Bruck, 2012), interviewers are strongly encouraged to resort to using open-ended and non-leading prompts in eliciting the most accurate information from children (Lamb et al., 2007). Under such circumstances where the structure of verbal interviews follows good-practice, do interview aids provide any added benefit? In Study 1, children's responses to *Wh*- questions were analyzed with and without the timeline regarding the target instance. Lamb, Hershkowitz, Sternberg, Esplin, et al. (1996) found the lengths of children's responses to invitational utterances (e.g., "*Tell me everything …*") and focused-recall questions (e.g., "*What color was …*") differ significantly, thus the effect of the timeline cannot be generalized across different question types. As such, Study 2 addressed this issue by comparing children's recall narratives of the target event in general (script memories) as well as a target instance (episodic memories) in response to invitational prompts, with or without the timeline as a guide. The event procedure followed that of Study 1. However, instead of asking focused-recall questions, children were asked to recall everything they remembered about the target activities (in general), and everything they remember about the last day of the activities (target instance). A few studies have shown that dolls are not associated with an increase in errors nor the amount of correct information produced in an interview that is consistent with the NICHD protocol, which encourages using open prompts (e.g., Salmon et al., 2012, Exp 2). As such, in Study 2, the timeline was not expected to be any more helpful than the verbal interview where only invitational prompts were used. These studies are novel in that the timeline played different roles and were tested with various types of questions. In Study 1, children used the timeline as a guide

to recall specific details regarding the target instance and as a tool to demonstrate their temporal memory for event details. In Study 2, children's spontaneous reports of target details for a specific instance were guided by the presence of a timeline where no interaction with the tool was needed. The challenges that children and interviewers face in providing and retrieving instance-specific memories have been well documented in empirical research; forensic interviewers are in dire need of developmentally-appropriate techniques that could help children report episodically. Findings from these two studies could shed light on whether the timeline could be an appropriate tool for children in retrieving episodic memories that are distinct from a general event representation.

A full assessment of the timeline in the forensic context not only includes examining children's recall accuracy using the tool, but also an investigation into how credible timelines are perceived to be and how well children seem to use them. Specifically, adults are the ones that determine whether children's testimony is accurate (when no physical evidence is available), as a result, the perceived accuracy should be as important as the actual accuracy (Castelli, Goodman, & Ghetti, 2005). No published work has examined whether children's perceived credibility is affected by whether and how they use interview aids. To address this gap in the literature, Study 3 examined whether the role of the timeline, from an adult's perspective, differs from what we have gathered from research showing that the timeline as an interview aid may or may not be suitable for child interviews.

Chapter 2

Study 1: The timeline: Recall and source monitoring of a specific instance of a repeated event

Using an experimental design, the present study investigated whether timelines helped 6- to 9-year-old children recall memories from specific instances of a repeated event. Children who participated in the current study were recruited from Taipei, Taiwan as part of a larger sample that was looking at cultural differences in children's narrative recall. Given the fundamental processes of the variables examined, culture differences were not a consideration.

The repeated-event experimental design ensures the standardization of procedure, the control of target stimuli, as well as the evaluation of accuracy. The Gosse and Roberts (2013) study added to our understanding of children's use of a timeline to aid the retrieval of events that happened within a day (i.e., day-scale), but child sexual abuse complainants often allege protracted repeated offenses. Thus, it is important that we extend our knowledge of whether timelines may aid children in identifying temporal information for repeated events that stretch over a longer time period. As such, we used a time scale to depict events that occurred over a two-week span.

Hence, we asked two main questions: 1) Do timelines help children recall one specific instance of a repeated event? and 2) Do timelines help children make accurate source decisions regarding when specific details occurred? Past visual aid research often asked children to demonstrate with and respond to questions using a prop, and have led to quantity-accuracy tradeoffs. In addition, we explored whether the timeline would pose risks of suggestibility to children's recall. In the current study, the timeline was evaluated under two different circumstances. First, when eliciting children's memories of the target event through recall, the

timeline was used as a reminder to children about the different occurrences, rather than as something that children must interact with symbolically. This may reduce memory demands and yet it does not require concurrent demonstration of child-to-symbol mapping. Second, children interacted with the timeline when later asked to show the source (i.e., the occurrence) of particular details. Findings could provide further explanation of how, under what circumstances, and for whom timelines do and do not work.

As previous research has consistently shown that younger children have difficulties with understanding objects as representations of other entities (e.g., DeLoache & Marzolf, 1995), timelines may best serve older children who at least have a basic understanding of time. Thus, we hypothesized as follows:

1) Older children would be more accurate and make fewer source-monitoring errors than their younger peers at both verbal recall and when using a timeline.
2) Children's accuracy would differ between verbal (control) and timeline conditions. Although a positive effect of the timeline would be desirable from a practical point of view, it is premature to suggest the direction of the effect due to the mixed findings and limited research on this topic.

Method

Design. The study employed a 2 (Age: younger, older) x 2 (Interview condition: Control, Timeline) between-subjects design. All children first participated in the Laurier Activities (hereafter "the Activities") on four occasions, and were later interviewed twice – immediately after the event when misleading information was presented; and a week later. Children were randomly assigned to either the control or the timeline interview condition balanced across gender and age.

Participants. One hundred and seventy-one children whose native language was Mandarin were recruited from elementary schools in New Taipei City, Taiwan. Forty-six were excluded because they missed an event or interview, and 4 were excluded due to interviewer error. The final sample comprised 121 children (53 boys): 60 children in the younger age group (M = 7.40 years-old, SD = 0.43) and 61 children in the older age group (M = 8.91 years-old, SD = 0.42). See Table 3 for mean ages in each Age x Interview Condition cell.

Ethical approval was obtained through the relevant institutional review boards. Parents provided informed consent and children gave verbal assent prior to each event and interview. Children were compensated with a small gift bag, and parents received a gift card (value of USD $3.5).

Materials. All materials used in the event and interviews were in Mandarin, and the events and interviews were conducted by native speakers. Materials, sample quotations, and the content of the interview protocols described in this manuscript have been translated to English for dissemination.

*The repeated event***.** The four occurrences of the Activities were identical in structure, lasted approximately 15 to 20 minutes, and contained 16 target items that centered around five main activities (i.e., puzzle, stickers, free drawing, connect-the-dots, and a story). Of the 16 target items, six were fixed in that the item remained the same across occurrences (e.g., the leader wore a red cloak every time) and the remaining 10 items were variable in that the instantiation, which is a version of the detail, was different every time. For example, children always completed a puzzle but the puzzle was different each time (i.e., four instantiations). On the target (last) occurrence, each child received a candy badge.

To control for item effects, two sets were created (see Table 2). Set 1 consisted of instantiation groups A, B, C, and D; set 2, groups D, B, C, and E. The design of the event and target details followed that of Powell et al. (1999) and Brubacher et al. (2011).

The biasing interview. Sixteen suggestive questions about the target occurrence (candy badge time) were administered, one for each target item. Two counterbalanced versions of suggestions were used (see Table 1 for items and their suggestion types). For the fixed items, each version included three items with a false suggestion that children had never experienced or seen in the Activities; and three items with a true suggestion. The true suggestions were included so that the children would not grow suspicious of the misleading questions.

A false suggestion was present in each of the 10 variable items. Five contained a false-internal suggestion, selected from experienced instantiations from non-target occurrences (e.g., for the puzzle on the last day of the Activities, it would be suggested that it was a snowman puzzle when the snowman was actually assembled on the second day of the Activities). The remaining 5 items contained a false-external suggestion created from non-experienced instantiations (i.e., from column E and A for item sets 1 and 2, respectively), such as suggesting that there was a beach ball puzzle on the last day when there never was a beach ball puzzle in any occurrence. For example, "*What was the color of the beach ball puzzle, the time you got a candy badge at the Laurier Activities*?" was asked when a child had not experienced the beach ball puzzle at the Laurier Activities. All suggestive questions used the label "the time with the candy badge" to make clear that the questions were referring to the target occurrence.

The recall interview. Sixteen focused-recall questions (one for each target item) were used to probe children's memory of the last event (e.g., *"What was the puzzle the time you got a candy badge?"*).

The 10 instance attribution questions tested whether children could attribute individual instantiations to their respective occurrence. There was one question for each variable item in the Activities. Fixed items could not be included here because it would not be possible to test which occurrence the child was recalling as fixed items were identical in each occurrence.

Two target items from each of the first, second, and third occurrences, and four target items from the fourth occurrence were randomly selected. Half were items that received a false-internal suggestion in the biasing interview; and the other half were items that received a false-external suggestion. See Figure 1 for interview structure and suggestion types.

The timeline. A horizontal straight line was drawn on a letter-sized sheet of paper with four anchors equally spaced on the line. The anchors were labeled with the numbers 1 through 4 to depict the four instances of the Activities (See Appendix A). The title of the Activities was printed horizontally on the top. In addition, the candy badge that children received during the last occurrence was placed underneath '4' to indicate that it was the time they received the candy badge.

Procedure. *The repeated event.* Children participated in the Activities on four occurrences over a two-week span. Two occurrences took place in the first week, separated by 2 days; the other two occurrences, the following week. Each occurrence was led by the same female research assistant, known as the Leader, and all occurrences were held in the same room in the children's schools. A script was followed during the Activities and the verbal label and visual presentation of each target item were properly and consistently presented across the Activities. A candy badge was given to each child at the beginning of the last occurrence as a way of referring to the target occurrence during the interviews that followed.

The biasing interview. This first interview was held three to five days after the target (last) occurrence and is referred to as the biasing interview because false information was presented. Children were individually interviewed by an unfamiliar female research assistant for 15-20 minutes in a different room to where the Activities were held. After a brief rapport-building session, the interviewer asked the child to recall everything they remembered about the Activities (in general). Then, 16 suggestive questions about the target occurrence were asked, one for each target item.

Prior to discussing the target occurrence, children in the timeline condition were presented with the timeline and the candy badge and instructed to refer to them in answering questions about the time with the candy badge. Children did not interact with the timeline. Children in the control (no timeline) condition were presented with only the candy badge and reminded that these questions were about the time they received the candy badge. The suggestive questions were then asked in random order. The purpose of these questions was to suggest instantiations that may not have occurred during the target occurrence. Past research has shown reliable suggestibility effects using this type of pre-suppositional questioning (e.g., Powell et al., 1999; Roberts et al., 2000). In other words, children demonstrate suggestible behaviors by accepting the false information planted in the questions.

The recall interview. The purpose of the second interview was to elicit children's memory of the target event through 16 focused-recall questions, one for each target item. A week after the biasing interview, children were non-suggestively interviewed by the same interviewer from the biasing interview. The timeline procedure and candy badge reference were presented as in the biasing interview.

The 16 focused-recall questions about the target occurrence were then asked. Before asking the 10 instance attribution questions, children were told: *"For the following questions, I need you to tell me which time at the Activities something happened. You need to tell me whether it was the first time, the second time, the third time, or the time with the candy badge"*. Children in the timeline condition were asked to respond by pointing to the instance number on the timeline, while those in the control condition received only the verbal instructions. The list of occurrences (from 1st to 'candy badge time') were emphasized throughout the interview.

Coding of the recall interview *Memory questions.* The coding system for the responses to the memory questions was based on those used in the Powell and Roberts (1999) study. All categories are mutually exclusive. See Table 4 for the codes and definitions for the responses from children's recall of the target (last) occurrence (the biasing interview was not included as it was merely a vehicle to deliver the suggestions).

Instance attribution questions. Children's responses to the instance attribution questions were coded as correct, incorrect, or 'I don't know/remember'. Other responses included multiple occurrences (e.g., "The first or the second time"), denial that such an item was present, or claims that 'it was every time' were rare and not analyzed further.

All transcripts were coded by a trained research assistant and, throughout coding, the author coded 15% of the transcripts from each Interview Condition x Age group. The small number of disagreements were resolved through discussion. Interrater reliability (kappa) for the coding of biasing, memory, and instance questions was .94, .92, and .97, respectively.

Results

Significance level was set at $\alpha = .05$ and Bonferroni corrections were applied when appropriate. All within-subject dependent variables in Analyses of Variance (ANOVAs) are described as proportions. The proportions were calculated by dividing the number of responses in each category (e.g., accurate responses, intrusion errors) by three for fixed items (3 true and 3 false suggestions), and by five for variable items (i.e., 5 false-external and 5 false-internal suggestions). For example, the accuracy for fixed items with a false suggestion was calculated by dividing the number of accurate responses for fixed items with a false suggestion by 3. Similarly, the accuracy for variable items with a false-internal suggestion was calculated by dividing the number of accurate responses for variable items with a false-internal suggestion by 5. See Table 1 for items and their corresponding suggestion type in the biasing interview. 'Suggestion Type' refers to how items were described in the Biasing Interview.

Accurate reports of the target (last) occurrence

Separate 2 (Interview Condition: Control, Timeline) x 2 (Age Group: Younger, Older) x 2 (Suggestion type: True, False or False-internal, False external) repeated-measures ANOVAs with the last variable within-subject were conducted to examine whether children's accuracy for the target items differed by their interview condition (H1) and/or by age group (H2). Suggestion type for fixed items consisted of true and false suggestions; and for variable items consisted of false-internal and false-external suggestions.

Fixed items. Overall, children's accuracy for fixed items were close to ceiling ($M = .96$, $SD = .09$). A significant effect of interview condition, $F(1, 117) = 5.21$, $p = .02$, $\eta_{p2} = .04$, which was subsumed by a significant interaction between suggestion type and interview condition, $F(1, 117) = 9.91$, $p = .002$, $\eta_{p2} = .08$. When using a timeline only, children were more accurate for

items that were not falsely suggested than for those that were falsely suggested in the biasing interview (M_{true} = .97, *SD* = .09; M_{false} = .89, *SD* = .19), as would be predicted in a suggestibility paradigm, $t(60) = 2.96$, $p = .004$, 95% CI [.03, 0.14], Cohen's $d = 0.55$. No difference was found in the control condition ($p = .17$). Further results from independent samples *t*-tests revealed that the interview condition effect was specific to items that were falsely suggested, $t(77.74) = 3.55$, $p = .001$, 95% CI [0.04, 0.14], Cohen's $d = 0.64$. Children in the control condition were more accurate than those in the timeline condition to questions about falsely-suggested items ($M_{control}$ = .98, *SD* = .07; $M_{timeline}$ = .89, *SD* = 0.19).

Variable items. The repeated-measures ANOVA yielded a significant effect of age group, $F(1, 117) = 6.70$, $p = .01$, $\eta_{p2} = .05$. Older children were more accurate than younger children (M_{older} = .43, *SD* = .17; $M_{younger}$ = .35, *SD* = .17). No other differences were found (see Table 5 for the full set of means).

Reports of suggestions from the biasing interview

To explore the question of whether the timeline posed risks of suggestibility, the following analyses were conducted.

Fixed items. A floor effect was seen in children's reporting of false suggestions for fixed items ($M = 0.07$, $SD = 0.31$), so 2 (Interview Condition or Age Group) x 2 (Reported at least one suggestion: Yes, No) Chi-square tests were used to analyze the associations between the likelihood of reporting a suggestion for the fixed items and interview condition and age group, respectively. Results revealed no statistical significance for either test, χ_2 $(1, N = 121) \leq 2.74$, $ps \geq .10$.

Variable items. A 2 (Interview Condition: Control, Timeline) x 2 (Age Group: Younger, Older) x 2 (Suggestion Type: False-internal, False-external) repeated-measures ANOVA revealed a significant effect of suggestion type, $F(1, 117) = 98.23$, $p < .001$, $\eta_{p2} = .46$. Children reported the false-internal suggestions from the biasing interview more often than the false-external suggestions ($M_{false\text{-}internal} = .26$, $SD = .22$; $M_{false\text{-}external} = .05$, $SD = 0.09$).

There was also a significant interaction between age group and interview condition, $F(1, 117) = 4.45$, $p = .04$, $\eta_{p2} = .04$. Simple effect analysis using independent *t*-tests revealed that older children reported proportionally more suggestions when using the timeline than same-age controls, ($M_{control} = .12$, $SD = .10$; $M_{timeline} = .18$, $SD = .12$), $t(59) = 2.11$, $p = .04$, 95% CI[-.12, .003] , Cohen's $d = 0.55$; there was no effect of interview condition among younger children, $p = .38$. Means show that younger children in the control condition reported more suggestions than their older counterparts, although the two-tailed significance was marginal at best, $t(58) = 1.76$, $p = .08$, 95% CI[-.01, .11], Cohen's $d = 0.46$. There was no effect of age group within the timeline condition, $t(59) = -1.24$, $p = .22$. Reports of false-external suggestions were rare, yet 2 (Interview Condition: Control, Timeline) x 2 (Reported at least one suggested detail) Chi-square tests also revealed no differences between interview condition, χ^2 (1, $N = 121$) = 2.19, $p = .14$), nor age group, χ^2 (1, $N = 121$) = 0.37, $p = .54$ (see Table 6 for the full set of means).

Source Confusions

In the above analyses, it cannot be determined which of these errors were due to the suggestibility process versus spontaneous source confusion between the instances. Thus, we also examined internal intrusion errors (reporting non-target instantiations as if they were present in the target instantiation) independent of the times that children reported suggestions. These

analyses of intrusion errors tested the difference between the errors made by younger and older children and whether the timeline had an effect on children's erroneous reports (H2).

Internal intrusion errors (variable items only). A 2 (Interview Condition: Control, Timeline) x 2 (Age Group: Younger, Older) x 2 (Suggestion Type: False-internal, False-external) repeated-measures ANOVA was used to analyze the proportions. A main effect of age group was found, $F(1, 117) = 4.00$, $p = .05$, $\eta_{p2} = 0.03$. Older children spontaneously made proportionally fewer internal intrusion errors than younger children ($M_{older} = .26$, $SD = .15$; $M_{younger} = .32$, $SD = .17$). There was a significant main effect of suggestion type, $F(1, 117) = 53.92$ $p < .001$, $\eta_{p2} = .31$. Spontaneous internal intrusions regarding items where false-external suggestions were presented (at the biasing interview) were double those with a false-internal suggestion ($M_{false\text{-}internal} = .19$, $SD = .16$; $M_{false\text{-}external} = .38$, $SD = .25$; see Table 7 for the full set of means).

Of all internal intrusion errors, 19% and 28% came from the first and second occurrences, respectively. The remaining 53% came from the third occurrence (i.e., the one closest to the target occurrence), which demonstrates a recency effect. To further examine the characteristics of the internal intrusion errors and whether the timeline had any influence on those errors, the distance of all internal intrusion errors from the correct (target) occurrence was calculated by using the absolute value of the difference between the target occurrence (i.e., 4) and the reported occurrence (i.e., 1, 2, or 3). An exploratory two-way ANOVA by age group and interview condition revealed a significant main effect of interview condition, $F(1, 109) = 6.61$, $p = .01$, $\eta_{p2} = .06$. Children's internal intrusion errors were closer to the target occurrence in the timeline condition versus controls ($M_{control} = 1.79$, $SD = 0.54$; $M_{timeline} = 1.55$, $SD = 0.50$; see Table 8 for the full set of means).

Instance attributions questions

The following analyses examined age differences (H1) and interview condition effects (H2) on children's accuracy in making source attributions for the variable items using a 2 (Interview Condition: Control, Timeline) x 2 (Age Group: Younger, Older) x 2 (Suggestion Type: False-internal, False-external) repeated-measures ANOVA with the last variable within-subjects.

Accuracy. Children of both age groups and interview conditions performed equally well regardless of suggestion type, Fs ≤ 1.81, ps ≥ .11, η_{p2}s ≤ .02). Children's accuracy for both false-internal and false-external items were comparable to chance. For items described at the biasing interview with a false-internal suggestion, the chance of matching the correct instantiation was .25 (matching to one of four instances); for items with a false-external suggestion, children's chance of matching the correct instantiation was lowered to .20 due to the additional option of the false-external suggestion. Results revealed that children matched instantiations to their respective occurrences significantly above chance for both *false-internal* and *false-external items*, ps < .002, with the exception that the accuracy for younger children in the control condition was not different to chance (p = .21).

Of the 10 instance attribution questions, 2 came from each non-target occurrence (i.e., Occurrences 1-3), and 4 came from the last (target) occurrence. Further exploratory analyses on the characteristics of the accurate and inaccurate attributions were conducted. First, a 2 (Interview Condition: Control, Timeline) x 2 (Age group: Younger, Older) x 3 (Occurrence: First, Middle two, Last) repeated-measures ANOVA on the proportion of correct attributions revealed a significant main effect of occurrence on accuracy, F(1.54, 180.39) = 51.74, p < .001, η_{p2} = .307. Paired-samples t-tests revealed that accurate instance attributions of the last occurrence (M = .77, SD = .53) were more accurate than that of the first occurrence (M = .33, SD = .34), t(120) = -7.55, p < .001, 95% CI[-0.55, -0.32], Cohen's d = 0.97 and middle two occurrences (M = .33, SD

= .21), $t(120) = -8.129$, $p < .001$, 95% CI[-0.55, -0.33], Cohen's $d = 1.08$). These effects did not differ by age group nor interview condition, Fs ≤ 1.47, $p \geq .477$, η_{p2}s $\leq .012$ (see Table 9).

Second, the distances between the occurrence erroneously stated by children and the correct occurrence were entered into a 2 (Interview Condition: Control, Timeline) x 2 (Age Group: Younger, Older) x 2 (Suggestion Type: False-internal, False-external) repeated-measures ANOVA and results revealed a significant effect of suggestion type, $F(1, 115) = 4.59$, $p = .03$, $\eta_{p2} = .04$. When items were misled with a false-internal suggestion at the biasing interview, children's erroneous instance attributions were closer to the correct instance than when a false-external suggestion was provided earlier ($M_{false\text{-}internal} = 1.47$, $SD_{false\text{-}external} = 0.44$; $M_{false\text{-}external} = 1.59$, $SD_{false\text{-}external} = 0.42$), regardless of age group or interview condition (see Table 9 for the full set of means).

I don't know/remember responses

Memory questions. Though we did not make predictions with regard to the proportion of 'I don't know/remember' responses, nearly 40% of the children responded so at least once during the memory questions about variable items. A two-way ANOVA revealed a significant interaction effect of age group and interview condition, $F(1, 117) = 8.11$, $p = .005$, $\eta_{p2} = .07$. Age differences were significant only in the control condition such that older children provided more 'I don't know/remember' responses than their younger counterparts ($M_{older} = .14$, $SD = .19$; $M_{younger} = .05$, $SD = .10$), $t(41.66) = -2.44$, $p = .02$, 95% CI[-.18, -.02], Cohen's $d = 0.64$. Further, older children provided more of these responses in the control ($M = .14$, $SD = .19$) versus the timeline condition ($M = .05$, $SD = .08$), $t(37.76) = 2.39$, $p = .02$, 95% CI[.01, .17], Cohen's $d = 0.62$.

Instance attribution questions. Fewer than 25% of the children provided any 'I don't know/remember' responses in response to these questions. Separate Chi-square tests were run (first for age group and then for interview condition) on the times that children reported at least one 'I don't know/remember' but there were no significant results, χ_2 (1, N = 121) ≤ .07, ps ≥ .54.

Discussion

Timelines are often used as interview aids in investigations of child abuse to help children identify a single instance from multiple experiences of abuse (Gosse & Roberts, 2013). As studies on the role of this type of visual aid are scarce, the effects of children's use of a timeline were systematically examined in the current study. Mixed findings of positive and negative effects on recall and source monitoring were found using a repeated-event paradigm: 1) Older children performed slightly better when the timeline was not used; 2) the timeline increased suggestibility for items that remained constant throughout the Activities; 3) the timeline did not affect the accuracy of children's source-attribution responses. Our specific findings are now discussed from the perspectives of recall, source accuracy, suggestibility, and development, and how using the timeline influenced each measure.

Effects of the timeline

The timeline had an effect on children's recall of fixed items. Their responses to the memory questions showed that accurate and erroneous reports varied not only by interview condition, but also by the type of detail (fixed vs. variable) and the type of suggestion in the biasing interview (true vs. false for fixed items; false-internal vs. false-external for variable items).

Regarding fixed items, children who used a timeline showed a reliable suggestibility effect such that they were less accurate for the misleading false items (provided to them in the biasing interview) than for non-misleading items. The purpose of using a timeline is to give children a concrete visual representation of all the instances of the event so that they can make better source decisions. However, the nature of fixed items is that they are the same throughout the Activities, so the process of making an association for a fixed item with one particular occurrence is not necessary. As such, it is possible that the presence of the timeline confused children and increased the impact of false information about fixed items. Additionally, using the timeline may have distracted children from the main task at hand. This is an important finding, as children are usually highly accurate in recalling details that remain the same throughout a series of events (e.g., Brubacher et al., 2011; Powell et al., 1999). Indeed, children in the current study were highly accurate overall for fixed items, and the proportional difference in their scores was not large. However, the presence of the timeline did not help children resist accepting misleading information even for items that children would normally recall with high accuracy using minimal cognitive resources. Summarily, the finding that the addition of a visual aid increased errors for fixed items is alarming, especially when forensic investigators often do not know which child-reported details are of a fixed versus variable nature.

The inclusion of the timeline in the current study did not improve accuracy for items that varied. In contrast to information that is always the same, children often confuse details that vary across occurrences of a repeated event (e.g., Connolly & Lindsay, 2001; Powell et al., 1999). Children not only have to recall the content of the details (the *what*) but also to discriminate their temporal source (the *when*) (Brewer & Williams, 2005). Thus, asking children to retrieve the correct version of a detail from a target occurrence was as difficult for the children in this study

as has been demonstrated in other studies of repeated events (e.g., Brubacher et al., 2011; Connolly & Lindsay, 2001; Powell et al., 1999) even with a representational aid designed to reduce cognitive demands. This highlights the immense difficulty that children have when asked to particularize.

In order to determine whether timelines are appropriate for investigative interviews, examining the characteristics of children's erroneous reports in conjunction with the timeline is important. Although the timeline did not improve accuracy, its use was significantly associated with the number and type of errors reported. Different patterns of errors were seen, particularly with respect to internal intrusions. Children in the timeline condition intruded details from the occurrence closest to the target occurrence more often than those in the control condition. The same pattern was found in responses to the instance attribution questions. Thus, the timeline may be 'pushing' the occurrences closer together in children's mental representations of repeated events. It is not clear why the timeline would lead to such an outcome. On one hand, memories of the 3rd and 4th occurrences contained similar temporal information as they were presented close in time. This would lead to increased source errors (Lindsay, Johnson, & Kwon, 1991; Roberts & Blades, 1998). That is, the perceived shorter distance by those who used a timeline could have increased source errors about occurrences that are mentally represented as close together in time.

A contrasting explanation emerges when focusing on responses from children in the control condition. These children confused occurrences from across the series rather than just the one closest to the target occurrence. After repeated experiences, children just like adults tend to develop scripts – cognitive structures representing the gist of what usually happens (Hudson & Nelson, 1986). Although the order of events in a script can be recalled accurately, temporal

information that is unique to individual instances is typically not retained in a script. As children in the control condition intruded more details from earlier occurrences (vs. those who used a timeline) it suggests that the individual occurrences were being integrated into a script (i.e., although individual instantiations were remembered they were not tied to any particular occurrence). This is further supported by the finding that younger children showed evidence of intruding from the 3rd occurrence than did older children. Younger children need more repetitions of an event to form a script than do older children (Farrar & Goodman, 1992). Thus, the timeline may have functioned to disrupt or suppress retrieval by script. Further research is necessary in addressing these issues.

Developmental differences

Typical developmental trends were present in our study in that older children were generally more accurate and less suggestible than younger children (Bruck & Ceci, 1999), although there was little difference in responses from older and younger children when using the timeline. In addition, there was a pattern showing that younger children reported proportionally more suggestions than the older children when no timeline was used (consistent with traditional findings). Moreover, older children's more advanced developmental capabilities including working memory, intelligence, and language abilities may lead to superior reports versus those from younger children (e.g., Eisen, Goodman, Qin, Davis, & Crayton, 2007). Interestingly, among the older children, those who were exposed to the timeline reported a greater number of false suggestions (vs. controls). In line with the lowered accuracy associated with some other types of visual aids (e.g., Bruck et al., 1995), the pictorial representation of time did not help older children resist suggestions. Perhaps this reflects that older children did not need to use the timeline because they were already proficient in temporal memory (Friedman, 1993; Gosse &

Roberts, 2013), and that using the timeline merely distracted or confused them. The findings may also reflect a developmental reversal (Brainerd, Reyna, & Ceci, 2008; Connolly & Price, 2006) - older children would have identified the similarity and gist of the events sooner than the younger children. This does not explain, however, why the timeline itself led to a greater confusion as all of the older children would be able to connect the meanings between individual occurrences. Previous research has questioned whether older children, who indeed have a well-developed understanding for symbolic representation, may be too old to benefit from visual aids. Future research is necessary to determine whether a visual aid increases cognitive demand that could lead to such confusions for older children.

The type of previously provided misleading information had an effect on children's later recall. More errors were made when children were initially misled with a version of the detail they had not experienced at all (false-external), than when they were misled with an experienced detail (false-internal). This is a novel finding that gives insight into the power of misleading questions on children's memories of a repeated event. When a non-experienced suggestion was given, children later had an additional option to consider when attempting to retrieve memories for the target occurrence. This may have required more cognitive resources that are highly correlated with age (e.g., working memory). The external suggestions may also have caused retroactive interference and demanded additional source decisions, all of which could increase the chances of making internal intrusion errors (the chance of recalling the correct instantiation decreased from 1:4 to 1:5). However, the timeline did not help reduce this cognitive load. This adds to the wealth of well-established research that emphasize the risks that suggestive questions pose to children's memories (e.g., Ceci & Bruck, 1993; Goodman & Melinder, 2007) with or

without the use of visual aids. As such, suggested questions should be avoided in interviews involving child witnesses (Lamb et al., 2007).

Compared to the 7- to 8-year-olds in the Gosse and Roberts (2013) study who matched events that happened within a single day to their parents' estimates on a timeline (roughly within one hour of their parents' estimates), older children in the current study did not demonstrate such ability on a larger time scale. It may be that children aged 9 years and above can make use of a timeline on the scale of a day, but they are not ready to use a timeline depicting longer stretches (Friedman & Lyon, 2005). Further research could tease apart these potential explanations by studying children's understanding and use of a timeline without memory demands.

Limitations and future directions

The cognitive demand when using a timeline as a visual aid is an important issue that needs more in-depth study. Though one of the first laboratory studies that systematically examined the role of a timeline in the recall of one specific instance of a repeated event, our study had limitations that can be addressed in future studies. First, the visual representation of the four occurrences on the timeline did not reflect that a weekend was between the second and third day of the Activities. Distinguishing weekdays from weekends is accomplished relatively early (Friedman, 1991) and so may have been helpful. Second, the timeline in Gosse and Roberts' (2013) study was anchored with landmark events such as pictures of breakfast and going to bed at the beginning and end of the timeline, respectively, whereas anchors used in this study were only limited to the occurrence order and the candy badge indicating the last (target) occurrence. Although as visual reminder of the four occurrences of events, the timeline did not contextualize the occurrences in any way. It is possible that children need the context provided by pictorial markers of events (e.g., breakfast) to remind them of what the timeline represents instead of

abstract information like numbers. Third, the same research assistant provided the biasing information in the biasing interview as well as conducting the later memory interview. Due to a constraint in the scheduling of the research assistants involved in the study, this design limited our findings on whether children were genuinely misled or complying to the interviewer when they reported previously suggested false information in the memory interview.

Conclusions

This is the only study that we know of that has tested the use of a timeline when recalling repeated instances of an event following misinformation. The results were clear - children's ability to retrieve and attribute details to a particular instance was not improved by using a timeline. Thus, if a timeline is used, it is recommended that investigative interviewers exhaust children's verbal recall using open-ended and non-leading questions before introducing any visual aids because they may not provide any added benefit to a good-quality verbal interview (Lamb et al., 2007). Children's understanding of the meaning of a timeline also needs to be more clearly established, along with consideration of how to familiarize children with a representation of time (e.g., in a practice phase of an interview).

Most apparent is the need for systematic, developmental research to ascertain the temporal and memorial mechanisms needed to use timelines depicting various temporal circumstances (e.g., a day, a year). The risks and benefits of using timelines may fluctuate depending on the type of memory task (e.g., delay, recall, number of similar occurrences). Future research could focus on answering the questions of why, how, when, and with whom timelines can be used effectively, and under what other circumstances (c.f., those in the current study) timelines may be beneficial to child witnesses. Study 2 further explored these questions by examining the timeline in interviews that used open-ended questions.

Chapter 3

Study 2: The timeline: The specificity of spontaneous recall of a specific instance of a repeated event

There is strong evidence supporting the use of open-ended prompts in forensic interviews. Open-ended prompts (*"Tell me everything ..."*) provide children the opportunity to report from recall memory and have been shown to lead to the fewest recall errors compared to forced-choice questions (e.g., Yes/No questions)(e.g., Lamb & Fauchier, 2001). Based on evidence from empirical research, the NICHD protocol was developed to encourage child interviewers to use open-ended prompts to minimize inadvertent contamination of children's accounts. For example, Lamb et al. (1999) demonstrated that prompts, which invite open-ended recall, led to the lengthiest and the most informative responses from children compared to focused-recall questions, which focused on specific details.

Anatomical dolls and human figure diagrams have been studied for their effects in interviews where open-ended recall questions were asked. Brown et al. (2007, Exp 1) employed a design that conducted an interview protocol based on procedures recommended by the NICHD protocol first and then an interview with or without a human figure drawing. Specifically, in their study, 5- to 7-year-olds' verbal recall was exhausted using open-ended recall questions first and then were asked about previously disclosed touches during the target event with or without a human figure drawing. Findings suggest that the drawing did not increase the number of correct details reported in response to open questions but elicited more incorrect details. In addition, the overall accuracy of children's reports was the highest in the control condition where no drawing was presented. These findings are consistent with the overall discussion of interview aid literature where the use of nonverbal props may increase the total amount of information recalled

at the expense of accuracy. Salmon et al. (2012, Exp 2) also conducted an NICHD protocol-based interview with 5- to 7-year-olds about a target event, which was followed by a series of follow-up questions about previously reported touches using a human figure drawing or an anatomical doll. No differences were found between the amount of information reported with a doll or human figure drawing and verbal questions only. Since the interviews followed that of the NICHD protocol, children's accuracies in the doll and human figure drawing conditions were not compromised. In summary, previous research collectively demonstrates that in an interview where open-ended prompts are utilized, the addition of other aids may not be needed. Children could perform just as well or better verbally compared to using an interview aid. Study 2 explored the use of a timeline to determine whether this type of aid conferred an advantage not seen with previously tested aids.

In addition to understanding whether the timeline affects the amount of target information recalled, another focus of Study 2 was to examine whether the timeline would affect children's reporting of episodic details, also known as particularization. Research has documented children's difficulty in particularizing (e.g., Connolly & Lindsay, 2001; Powell et al., 1999). In child sexual abuse investigations, it is crucial that child victims isolate individual occurrences of recurring abuse to create more credible testimony (Connolly et al., 2008). The repeated-event paradigm allows for the analysis of the target details reported from specific instances when timelines are or are not present.

In Study 1, it was shown that the timeline did not provide any added benefit to children's responses to focused-recall questions and source decisions about a repeated event. It would be unfair to conclude that the timeline is not helpful to children's recall without investigating other question types that are typically used in investigative settings. As such, in Study 2, children's

spontaneous recall of target details to open-ended prompts, accompanied by the presence of the timeline as a guide (timeline condition) or not (control condition) was investigated.

A group of Canadian children matching the ages of those in Study 1 participated in Study 2 using the same design. Children were interviewed 3 to 5 days after the last occurrence of the Activities. The structured interview consisted of two main phases. The first phase of the interview was about the Activities in general. In the second phase, children were asked to only talk about the target occurrence of the Activities, namely the occurrence during which children received a candy badge. Both phases followed recommendations from the NICHD protocol and began with an introductory phase, followed by an invitational prompt *(*"*Tell me everything about* ...") and two opportunities to elaborate ("*Tell me more about* ..."). The focal analyses of their recall narratives was detail level, characterized by whether details are item-level (i.e., puzzles, connect-the-dots) or instantiation-level (i.e., snowman puzzle, leaf connect-the-dots); and the specificity of their recalled details, characterized by whether details are generic (i.e., details that always happened) or episodic (i.e., details that happened in a particular occurrence). The hypotheses were as follows:

1. Although in Study 1, the timeline did not lead to more accurate responses to focused-recall questions, the study's findings and those from other visual aid research suggest that a visual aid may not promote memory for information per se. Study 2 tested the exploratory hypothesis of whether the timeline would help children report more instantiation-level and episodic details compared to the control group in an open-recall setting.
2. Older children would report more instantiation-level details than younger children. Previous literature has shown that older children tend to be more capable of reporting

instantiation-level details from a repeated event than younger children (Brubacher et al., 2012)

3. Older children would report more episodic details than younger children. Evidence from previous literature suggests that memory for a specific instance deteriorates more quickly for younger children than for older children (Powell & Thomson, 1996). As such, episodic details from specific instances may be retained better by the older group.

4. With respect to analyses of detail level and specificity coding, it was expected that children would report more item-level than instantiation-level details under the detail level coding. Similarly, it was expected that children would report more generic than episodic details under the specificity coding. Research has demonstrated that children tend to develop stronger scripts for highly similar recurring experiences (Farrar & Goodman, 1992). As these scripts become stronger after repeated experience, it becomes easier to identify details that always happen using generic language (Pearse, Powell, & Thomson, 2003).

5. With respect to analyses of recall phase, it was expected that more item-level and generic details would be reported in the general recall phase than in the target recall phase; and more instantiation-level and episodic details would be reported in the target recall phase than in the general recall phase. These findings would be consistent with previous research that showed children tend to report more episodic details when the interviewer used episodic language (Schneider et al., 2015); and with findings from Brubacher et al., (2012) that examined the item/instantiation-level details in the breadth (general recall) and depth (target recall) interviews.

Method

Design and participants. The study employed a 2 (Interview Condition: Control, Timeline) x 2 (Age Group: Younger, Older) between-subjects design. One hundred and seventeen children aged between 5 and 10 were recruited from primary schools in the Waterloo region. A total of 46 children were excluded from the study because they missed at least one day of the event and were subsequently not interviewed. The final sample of 71 children consisted of 26 boys and 45 girls, with a mean age of 8.23 years (*SD* = 1.06). Table 10 shows the sample size and mean age for each interview condition by age group cell.

The materials and event procedure were the same as that of Study 1.

Interview procedure. Children were individually interviewed by an unfamiliar research assistant 3 to 5 days after the last occurrence of the Activities in a room that was different from the event room. The interview included an introduction and was followed by a substantive interview. The substantive interview was structured and consisted of two phases: general recall and target recall. See Figure 2 for the interview procedure.

General recall. Children were asked to report "*everything you can remember about the Laurier Activities*". They were given exactly two follow-up prompts to elaborate without any information about the Activities (e.g., "*Tell me more about the Laurier Activities.*" or "*What else can you tell me about the Laurier Activities?*").

Before proceeding to the target recall phase, children in the timeline condition were introduced to the timeline. The introduction procedure of the timeline was the same as that from Study 1. The interviewer explained that the numbers 1 through 4 on the timeline represented the first, second, third, and the last time, respectively, that they were at the Laurier Activities. The

candy badge they received at the last occurrence was placed on the number 4 to represent that it was the time that they received the badge. Different from Study 1, no interaction with the timeline was requested by the interviewer; children were simply using the timeline as a visual guide. For the control condition, children were given the candy badge and were instructed that they were about to talk about the time that they received the candy badge at the Laurier Activities.

Target recall. Children were asked to recall everything they remember about the target instance (i.e., "*Tell me everything you remember about the Laurier Activities when you got a candy badge*"). Similar to the general recall phase, children received two follow-up prompts to elaborate (e.g., "*What else do you remember ...* ", "*Tell me more about ...* "). However, each follow-up prompt emphasized that they were to recall about the time they got the candy badge (e.g., "*Tell me more about the time you got a candy badge at the Laurier Activities*").

Coding. Interview audio recordings were transcribed verbatim for coding purposes. Children's reports of target details from the general recall and target recall phases were coded for their specificity.

Detail level coding. In each phase of the interview, mention of target items and/or instantiations were coded in three ways. First, they were coded for whether a detail was at an item level (e.g., "*There were puzzles*"), or instantiation-level (e.g., "*a snowman puzzle*" *or* "*a snowman*").

Specificity coding. The specificity of the language used to describe an item-level or instantiation-level detail was first coded for any associated temporal markers. For example, "*one time the puzzle was a snowman*" would be coded as having an episodic temporal marker; "*We always do a puzzle*" would be coded as having a generic temporal marker ("always"); and *"we*

did a puzzle" as having no temporal marker. Then, the grammatical tense of the sentence associated with any item or instantiation was coded. Categories include simple past tense, simple present tense, conditional tenses such as could/would, and no tense. Specificity was based on the associated temporal marker and grammatical tense. For example, "*We did a puzzle*" would be coded as episodic; "*We do puzzles*" would be coded as generic. Repeated details were not coded unless the detail was mentioned at a different level of specificity than it was originally reported. In addition, sentences in the same conversation turn shared the same temporal marker and grammatical tense. For example, "*We did puzzles, connect-the-dots*" was coded as past tense for both *puzzle* and *connect-the-dots*. Similar studies have also employed this coding scheme to determine the specificity of reported target details (e.g., Brubacher et al., 2011). Finally, if the temporal marker and grammatical tense were contradictory, then the temporal marker was coded. For example, "*We always did a puzzle*" would be coded as generic based on its generic temporal marker, but not the grammatical tense (which would deem the detail episodic). The temporal marker associated with a target detail always took priority in coding than the grammatical tense. See Table 11 for the specificity coding system.

All transcripts were coded by a trained research assistant and, throughout coding, the first author coded 15% of the transcripts from each Interview Condition x Age group cell. The small number of disagreements were resolved through discussion. Interrater reliability (kappa) for the coding of item/instantiation, grammatical tense and temporal marker was .97 and .93, respectively.

Results

Preliminary Analyses

Preliminary analyses were conducted to eliminate any effect of extraneous variables on the main dependent measures. Separate sets of independent samples *t*-tests were conducted using item set and gender as independent variables. No effect of item set or gender was significant on any level of the independent variables, *t*s ≤ 1.81, $p \geq .08$. Outliers for each dependent measure were identified by using the +/- 3 standard deviation selection criteria. As a result, 6 data points were removed. See Table 12 for more details. Significance level was set at $\alpha = .05$ and Bonferroni corrections were applied when appropriate.

Distribution. Due to the small sample size ($N = 71$), tests of normality were conducted in an attempt to qualify the use of parametric tests. Tests of normality indicated that all dependent measures were positively skewed due to many children recording a score of zero for the number of episodic or generic details or the number of instantiation- or item-level details. Index scores were thus calculated in an attempt to meet the normality assumption for parametric tests. Two index scores were calculated to represent the difference between the number of episodic and generic details reported in each recall phase. A positive episodic-to-generic index score means that more episodic details were reported compared to generic details. A further two index scores were calculated to represent the difference between the number of instantiation-level details and item-level details; a positive instantiation-to-item index score represents that more instantiation-level details were reported compared to item-level details. Tests of normality for the index scores revealed that the episodic-to-generic index score in the target recall phase and the instantiation-to-item score in the general recall phase remained non-normal. Analyses using the index scores yielded results consistent with analyses using the original measures. Since Analysis of Variance (ANOVA) and other parametric tests have been shown to be robust to violations of normality

(e.g., Blanca et al., 2017), main analyses of the original non-transformed raw scores using parametric tests are presented.

Manipulation check. Since children were only introduced to the timeline in the target recall phase, no interview condition differences should emerge in the earlier general recall phase. Independent samples *t*-tests confirmed that no dependent measures differed by interview condition in the general recall phase, ts ≤ 1.637, ps $\geq .11$, Cohen's ds ≤ 0.39. As such, the interview condition factor was only included in the analyses of the target recall phase.

Main Analyses

The hypotheses were tested separately for the numbers of details reported under the detail level coding and the specificity coding. First, to test the effects of the timeline and age on the numbers of item- and instantiation-level details in the target recall phase (H1 & H2), an interview condition by age group Multivariate Analysis of Variance (MANOVA) was conducted on the numbers of item- and instantiation-level details. To test whether children would report more item-level details compared to instantiation-level details (H4), and whether item-level details would be recalled more often in the general recall phase and instantiation-level details would be recalled more often in the target recall phase (H5), a repeated-measures ANOVA with recall phase and detail level as the within-subject variables was conducted.

Second, to test the effects of the timeline and age on the numbers of episodic and generic details in the target recall phase (H1 & H3), an interview condition by age group MANOVA was conducted on the numbers of episodic and generic details. To test whether children would report more generic details compared to episodic details (H4), and whether generic details would be recalled more often in the general recall phase and episodic details would be recalled more often

in the target recall phase (H5), a repeated-measures ANOVA with recall phase and detail level as the within-subject variables was conducted.

Six fixed items were included as target items. Overall, spontaneous reports of fixed items were rare in both recall phases ($M_{general\ phase}$ = 0.89, $SD_{general\ phase}$ = 0.70; $M_{target\ phase}$ = 0.44, $SD_{target\ phase}$ = 0.58). Since the main purpose of this study was to test the effect of the timeline on reports of specific details from various occurrences, only reports of variable items ($M_{general\ phase}$ = 3.85, $SD_{general\ phase}$ = 1.80; $M_{target\ phase}$ = 1.51, $SD_{target\ phase}$ = 1.36) were included in all main analyses. Descriptive statistics for the number of details reported by interview condition, detail level (instantiation-level, item-level), and specificity (episodic, generic) are displayed in Tables 13 and 14.

Spontaneous recall of target details. To test the effects of the timeline and age (H1 & H2), the analysis of variable details reported in the target recall phase was conducted using a 2 (Interview Condition: Control, Timeline) x 2 (Age Group: Younger, Older) MANOVA on the numbers of item- and instantiation-level details. No significant effects were found, Fs $\leq$ 2.77, p $\geq$ .07, $\eta_{p2} \leq$ .08.

Age effects on the detail level of reported details were also tested for the variable details reported in the general recall phase using a MANOVA with age group as the between group factor on the numbers of item- and instantiation-level details. There was a significant main effect of age group on the combined dependent variables, $F(2, 68) = 3.07$, Wilk's λ =.917, p = .05, η_{p2} = .08. Follow-up independent samples t-tests for item- and instantiation-level details separately showed a significant effect of age for the number of instantiation-level details reported, $t(69) = -2.17$, p = .03 , Cohen's d = 0.51, but not for the number of item-level details, $t(69) = -0.58$, p = .56, Cohen's d = 0.14. Older children (M = 3.17, SD = 2.98) reported a significantly higher

number of instantiation-level details than their younger peers (M = 1.78, SD = 2.43). No other effects were significant, Fs ≤ 1.81, p ≥ .17, η_{p2} ≤ .05.

Additionally, to test for whether the numbers of details reported varied by detail level and/or recall phase (H4 & H5), a 2 (Recall Phase: General, Target) x 2 (Detail Level: Item, Instantiation) within-subject ANOVA was conducted and yielded a main effect of recall phase, $F(1, 69) = 27.63$, $p < .001$, $\eta_{p2} = .29$. However, this main effect was qualified by a recall phase by detail level interaction, $F(1, 69) = 31.88$, $p < .001$, $\eta_{p2} = .316$. Simple paired-samples t-tests were conducted to examine this interaction. Results indicated that the number of item-level details reported significantly differed from instantiation-level details for both the general and target recall phases, ts ≥ |3.16|, ps ≤ .002, Cohen's ds ≥ 0.59. Children reported a significantly higher number of item-level details (M = 3.84, SD = 1.80) than instantiation-level details in the general recall phase (M = 2.46, SD = 2.78); while the trend was in reverse for the target recall phase where the number of instantiation-level details reported was higher ($M_{instantiation}$ = 2.89, SD = 2.16, M_{item} = 1.51, SD = 1.37). In terms of comparison between recall phases, only the number of item-level details differed, $t(70) = -9.36$, $p < .001$, Cohen's d = 1.46. More item-level details were reported in the general recall phase (M = 3.84, SD = 1.80) than the target recall phase (M = 1.51, SD = 1.36).

Specificity of the reported target details. To test the effects of the timeline and age (H1 & H3), the analysis of variable details reported in the target recall phase was conducted using a 2 (Interview Condition: Control, Timeline) x 2 (Age Group: Younger, Older) MANOVA on the numbers of episodic and generic details. No significant effects were found, Fs ≤ 1.97, p ≥ .15, η_{p2} ≤ .05.

Age effects on the specificity of reported details were also tested for the variable details reported in the general recall phase using a MANOVA with age group as the between group factor on the numbers of episodic and generic details. A main effect of age was significant on the combined dependent variables, $F(2, 65) = 3.18$, Wilk's $\lambda = .91$, $p = .05$, $\eta_{p2} = .09$. Follow-up independent samples *t*-tests for episodic and generic details separately showed no significant effect of age for either variable, ts $\leq$ -1.42, $p \geq .16$, Cohen's $d \leq 0.34$.

Additionally, to test for whether children would report more episodic than generic details (H4 & H5), the specificity of details reported was first analyzed using a 2 (Recall Phase: General, Target) x 2 (Specificity: Episodic, Generic) within-subject ANOVA. A significant main effect of recall phase, $F(1,66) = 17.93$, $p < .001$, $\eta_{p2} = .21$, was qualified by a recall phase by specificity interaction, $F(1, 66) = 107.96$, $p < .001$, $\eta_{p2} = .62$. Simple paired-samples *t*-tests were run to examine this interaction. The numbers of episodic and generic details differed in both the general and target recall phases, ts $\geq |6.21|$, ps $< .001$, Cohen's ds ≥ 1.24. In the general recall phase, more details were reported in the generic form ($M = 4.35$, $SD = 2.44$) than in the episodic form ($M = 1.63$, $SD = 1.88$); a reverse trend was seen in the target recall phase where more details were reported in an episodic form ($M = 3.57$, $SD = 2.26$) than in a generic form ($M = 0.80$, $SD = 1.30$). Comparing between recall phases, the number of episodic and generic details also differed, ts $\geq |5.86|$, ps $< .001$, Cohen's ds ≥ 0.91. A greater number of episodic details were reported in the target recall phase ($M = 3.54$, $SD = 2.34$) than in the general recall phase ($M = 1.64$, $SD = 1.89$); and a greater number of generic details were reported in the general recall phase ($M = 4.32$, $SD = 2.46$) than in the target recall phase ($M = 0.81$, $SD = 1.31$).

Developmental differences. As age differences were not clear through the between-group analyses due to individual differences in spontaneous recall and the small age gap between

the two age groups, bivariate correlations between children's age in years and the main dependent measures were explored (See Table 14 for the Pearson correlations). Age was significantly associated with the number of instantiation-level details reported in both the general (r = .25, p = .04) and target recall phase (r = .27, p = .02). As age increases, children reported more instantiation-level details (e.g., *snowman* puzzle). In addition, age increase was associated with more episodic style reporting in the target recall phase (r = .30, p = .01), as well as generic reporting in the general recall phase (r = .28, p = .02).

These developmental trends that emerged from correlation analyses prompted the analyses of main dependent measures using linear regressions with interview condition (control vs. timeline) and age in years as predictors. A total of 8 linear regression analyses were conducted (see Table 15 for standardized coefficients and test statistics). Overall, interview condition was not significantly related to any of the dependent measures. Age in years significantly predicted the number of instantiation-level details reported in both the general and target recall phase, ps $\leq$ .02, generic reports in the general recall phase p = .02, as well as episodic reports in the target recall phase, p = .01.

Discussion

Study 2 extended our understanding of the timeline as an interview aid for school-aged children's recall of a repeated experience. Children's spontaneous recall of target details with and without the timeline was examined. Consistent with findings from Study 1, the timeline did not show any added benefit over verbal questioning alone. Specifically, the timeline did not lead to higher numbers of recalled details in terms of the specificity nor detail level. Additionally, children's age did not contribute to any differences between the control and the timeline condition. In other words, neither older nor younger children in the study benefited from the

presence of the timeline in their recall. Findings regarding the characteristics of children's recall of target details from the repeated event are in line with previous research using the same repeated-event paradigm. The implications of these findings are now discussed.

Effects of the timeline

Different from Study 1 (instance attribution question phase), children did not physically interact with the timeline in their recall process (i.e., pointing at the timeline); the timeline was simply present at the time of recall and was used to facilitate children's recall of the target instance (i.e., details recalled in the target recall phase). We examined the effect of the timeline in two ways. First, we focused on whether the timeline would help children recall a fuller picture of the Activities by analyzing the detail level children's reports (general item vs. a specific instantiation of the item). Second, we focused on whether the timeline would help children recall more details about one specific instance of the Activities by analyzing the specificity of recalled details (generic vs. episodic detail).

As previously mentioned, a timeline is sometimes used by police and social workers as a way of eliciting temporal details or details that refer to a specific instance. In the current study, the focus was on episodic details that referred to a specific occurrence in the repeated experience (e.g., *"one time, the puzzle was a snowman").* Findings suggest that the timeline did not lead to an increase in the number of episodic details recalled in the target recall phase. Children who were visually presented with the timeline that showed all occurrences within the repeated event (numbered 1 through 4) as well as the occurrence of interest (numbered 4 and marked with the candy badge received for that occurrence), did not benefit from this visual cue. In addition, the number of generic details provided in the timeline condition did not differ from the control. Together, these findings are consistent with that of human figure diagram research, which

demonstrated that children who were interviewed without a human figure diagram reported as many details as those who used one when providing spontaneous recall about a target event (e.g., Salmon et al., 2012).

In the analyses of item-/instantiation-level details reported in the target recall phase, the effect of interview condition was not significant. Results from the linear regression analyses also did not show interview condition as a significant predictor of the number of item-/instantiation-level details. In other words, despite results showing that children in the timeline condition had an overall higher number of reported details, the timeline did not help children recall more item-level nor instantiation-level details about the Activities. Using Fuzzy-trace theory (Brainerd & Reyna, 2004) to explain the relationship between item- and instantiation-level details, more item-level details reported could represent that the children had stronger gist traces about the repeated experience; and more instantiation-level details reported could mean that the children retained stronger verbatim traces. Brubacher et al. (2012) suggested that by engaging in gist processing, which was asking children to recall generally about the Laurier Activities, the processing of individual verbatim traces (target recall) could be weakened. This could explain why in the current study, children recalled significantly more item-level (gist traces) in the general than in the target recall phase. In the target recall phase, children's recall could have been exhausted by their recall in the general phase. Thus, any potential difference between the control and timeline conditions could have been minimized. The timeline intended to serve as a visual reminder that there were four occurrences in total and that they were asked to only talk about the target occurrence. Having the visual cue available to distinguish between occurrences did not lead to children providing a fuller picture of what happened in the target occurrence.

Spontaneous recall of details from the repeated event

Despite the current study showing no support for the use of the timeline in recalling specific details (i.e., instantiation-level and episodic details) from a repeated event, results contribute significantly to our understanding of children's repeated-event memory. Repeated-event research in laboratory settings has advanced our understanding of its uniqueness compared to the memory of single events (e.g., Powell & Roberts, 1999; Brubacher et al., 2011; Connolly & Price, 2006). However, only a few studies have examined recall narrative ("*Tell me everything about* ...") rather than through a series of focused-recall questions (*Wh*- questions). The characteristics of children's recall of the repeated event in the current study are consistent with findings from a study by Brubacher et al. (2012) which shares a similar methodology. In the current study, children's spontaneous recall of the repeated event was examined based on the detail level (item-level vs. instantiation-level) and specificity (episodic vs. generic) of their recall. First, more item-level details were reported in the general than in the target recall phase suggesting that children may have exhausted their memory about the target event in the first recall opportunity. Similar to findings by Brubacher et al. (2012), the 7- to 8-year-olds reported significantly more item-level details in the breadth (general recall) phase than in the depth (target recall) recall. Second, an increase in the number of episodic details when only asked to give a broad recall of what happened over the repeated experience was associated with an increase in age. This is also consistent with findings from Brubacher et al. (2012) that also found a significant difference between the report of instantiation-level details from 4- to 5-year-olds and 7- to 8-year-olds. Further, a higher level of instantiation-level reporting by older children suggests that age significantly contributes to the ability to recall more specifically as evident in research using narrative or focused-questions formats (e.g., Brubacher et al., 2011; Brubacher et al., 2012).

Findings from the recall phase by detail level analyses demonstrated that the detail level and specificity of children's responses tend to correspond to that of the interviewer's. Specifically, when only asked to recall everything they remembered about the Laurier Activities, more item-level details were reported compared to instantiation-level details. The specificity of their reports also revealed that more generic details were reported compared to episodic details in the general recall phase where children were asked to recall everything about the repeated experience with no specific prompt about which occurrence. However, when asked about the target occurrence, children's responses shifted. Specifically, more instantiation-level details were reported compared to item-level details; the specificity of children's reports also shifted to be more episodic than generic. Together these findings suggest that children aged between 6- and 9-years-old are capable of recognizing the interviewer's style of questioning and adjusting their responses to match the questioning style. Schneider et al. (2015) also suggested that when interviewers asked a generic question to 4- to 16-year-olds, they received a generic answer; when interviewers asked an episodic question, they received a response back in an episodic language. In line with previous research (Schneider et al., 2015), findings from Study 2 suggest that in forensic interview settings, where the particularization of specific instances of abuse could substantiate child abuse allegations, an episodic questioning style that asks for details of a specific instance could elicit more instance-specific details than a general questioning style.

Developmental differences

In addition to the age effects discussed above with regard to recall phases and the type of details reported, this section focuses particularly on developmental differences with regard to recalling a repeated experience. Extant literature has focused on children's ability to remember a script (Hudson et al., 1992). Details that always occur in a repeated experience form a scripted

memory; details that deviate from the script require more cognitive resources to compare against the existing script (Farrar & Goodman, 1992). Literature shows that as children age, their ability to monitor similar sources (e.g., individual episodes act as different sources) is enhanced (Roberts, 2002). In the current study, the ability to distinguish separate episodes of the Activities was manifested by older children's ability to report more specific details. The focus of the analyses of detail level and specificity was on instantiation-level and episodic reports. These two detail types represented children's ability to recall more specific details (instantiation-level = more in-depth details; and episodic details = more about one episode rather than about the script). Age was a significant factor in children's reporting of episodic and instantiation-level details. As children's age increased, they reported a higher number of instantiation-level details in both recall phases regardless of the interview condition. This is consistent with previous work that shows that older children tend to report more details about the repeated experience than younger children regardless of the style of questioning (e.g., Powell & Thomson, 2000). Second, an age increase was associated with an increase in generic details reported in the general recall phase; and with an increase in episodic details reported in the target recall phase. The detail level analyses focused on the depth of information about the repeated experience children were able to describe, whereas the specificity of their report focused on whether they were able to describe a specific episode. As such, this shows further evidence that children's ability to report target details increases with age regardless of the interviewer's questioning style.

Limitations and future directions

One limitation of the current study, as also discussed in Study 1, was the visual presentation of the timeline used in the studies. The timeline did not reflect that there was a weekend in between the second and the third occurrence of the Activities and that the anchors

were simply numbers rather than pictorial markers of events that could help children differentiate between occurrences. Another limitation was that, by asking children to use the timeline as a guide without physically interacting with it, we could not ensure that children were actually processing the presence of the timeline while engaging in recall. Moreover, the amount of cognitive resources taken (if any) to process the timeline at recall was unknown. Future research should examine whether the presence of the timeline would occupy any cognitive resources and how that would affect children's recall and whether such cost of cognitive demand of using a visual aid poses additional risks to recall.

Conclusions

Taking the findings from Study 1 and 2 together, the evidence supporting the use of the timeline is weak in memory retrieval settings. When used as a facilitator to guide children's recall of a specific instance, no added benefit in terms of accuracy (Study 1) nor amount (Study 2) of recalled information was observed (i.e., analyses of instantiation-level and episodic details). In addition, by giving children a chance to respond nonverbally (i.e., pointing to the timeline) about which detail occurred during which instance of the repeated event (Study 1), source attribution accuracy did not differ compared to verbal responses. Findings from these two studies prompted a further question: if no evidence suggests that the timeline leads to any benefit in the recall of a repeated event, do adults hold an erroneous belief that a child's testimony that involved the use of a timeline can be more trusted? It is important that this question is addressed because, in legal processes, the credibility of the testimony of a child perceived by adult decision-makers could lead to major legal consequences. Study 3 addressed this question using a perceived credibility paradigm similar to that of Connolly et al., (2008).

Chapter 4

Study 3: Adults' perceptions of the timeline as an interview aid

In child sexual abuse investigations, child victims may provide testimonies that trigger further legal actions against alleged perpetrators. These actions could lead to successful convictions and/or involve irreversible consequences. The impact that these testimonies have on individuals involved in the legal cases are significant. As such, it is critical that children's testimony is accurate and credible. The credibility of children's testimony is determined by adults such as investigators, social workers, and legal workers, and is crucial to making legal decisions in how to proceed with their cases (e.g., Goodman, Batterman-Faunce, Schaaf, & Kenney, 2002). While the actual accuracy of children's reports is important in establishing a criminal investigation at an early stage, the credibility of their account becomes increasingly relevant as the investigation moves further in the legal process.

Adults' perceptions of child witnesses

In the last few decades, forensic psychologists have studied factors that impact adult jurors' decisions (e.g., Connolly et al., 2008). Empirical evidence from mock juror studies has yielded several factors that influence how adults perceive children's eyewitness reports. First, adults' perceptions of the credibility of child witness accounts differ from those of adolescents and adults (e.g., Cleveland, Quas, & Lyon, 2016; Peterson, 1996; Pozzulo, Lemieux, Wells, & McCuaig, 2006). Older children are often judged to display higher cognitive competence than younger children (Connolly et al., 2008). Second, children's confidence level, as perceived by adults, is positively associated with credibility (Cleveland et al., 2016; Schmidt & Brigham, 1996). Third, the consistency of testimony is considered to be an important factor used in determining the credibility of children's testimony by both actual and mock jurors (Granhag &

Strömwall, 2000; Myers, Redlich, Goodman, Prizmich, & Imwinkelried, 1999). Finally, other behavioral and contextual aspects such as the speed of identification and the use of visual aids affect adults' perceptions of eyewitness credibility (e.g., Neal, Christiansen, Bornstein, & Robicheaux, 2012). The current study examined whether and how adults consider these factors when determining children's credibility in recalling individual instances of a repeated event using a timeline as an interview aid. However, to the author's knowledge, no published work has examined whether using timelines (or any other visual aids) in child interviews influences the perceived credibility of children's recall from adults' perspectives. As a result, the following sections focus on reviewing literature that examined factors influencing adults' perceptions of children's credibility.

Age and cognitive competence. Evidence suggests that adults acknowledge that cognitive competence, such as memory capability (Quas, Thompson, Alison, & Stewart, 2005), differs significantly by age, and take this into consideration when deciding witness credibility. Some research suggests that children are sometimes perceived as less credible than adults (Leippe, Manion, & Romanczyk, 1993; Newcombe & Bransgrove, 2007), some studies found the contrary (e.g., Peterson, 1996; Ross, Miller, & Moran, 1987); and some studies found adults to be insensitive to age when deciding witness credibility (Cleveland et al., 2016). In a study that used mock trial transcripts of eyewitnesses describing events of a crime, a 20-year-old witness was rated as having significantly higher overall accuracy and reliability than 4- and 12-year-old witnesses (Bruer & Pozzulo, 2014). Similarly, Pozzulo et al. (2006) found that adults were perceived as more credible than 9-year-olds in their accounts of a car theft that led to a bank robbery. Findings from these studies suggest that when memory capability is of greater importance (i.e., when making identifications), adults may be seen to be more credible than

young children due to their advanced cognitive abilities (e.g., Bottoms, 1993). In contrast, an earlier study by Ross, Dunning, Toglia, and Ceci (1990) found that an 8-year-old child witness was rated as significantly more accurate, confident, honest, and credible than a 21-year-old witness in a mock trial case that involved a drug charge. Later studies confirmed that while younger children may be seen as less cognitively competent, their honesty could be valued over their cognitive abilities (Connolly, Price, & Gordon, 2010), as shown by in Ross et al. (1990) where a drug-related crime scenario was presented and honesty may be seen as more salient.

Several studies have also compared the perceived credibility of children of different ages and have yielded inconsistent findings. For example, 6- to 7-year-old children were judged to be more cognitively competent and credible than 4- to 5-year-olds when recalling details of a staged event (Connolly et al., 2008, Exp 1). A reverse age effect was found by Bottoms and Goodman (1994), where the testimony of a 6-year-old victim of sexual assault was judged to be more credible than that of a 14- and 22-year-old. The authors suggested that adults may have viewed the 6-year-old victim in the study to be less capable of fabricating details of a sexual abuse allegation, rather than considered their memory capabilities, thus rating them as more credible than the 22-year-old (Bottoms & Goodman, Exp 1, 1994). Together these findings suggest that, depending on the circumstances of the crime or methodology used in empirical studies, age may have a significant positive or negative influence on jurors' decisions.

Furthermore, adults have been found to be insensitive to developmental differences when controlling for other credibility factors. Cleveland et al. (2016) showed that adults did not differentiate 6- from 11-year-old witnesses in their perceptions of the credibility of children's temporal memory. However, when confidence level was considered, younger children who were confident were perceived as most accurate. In addition, the authors suggested that adults were

unaware of older children's ability to express uncertainty to adults' questions (e.g., "*I don't know*" responses); and this negatively affected children's credibility. Another study showed that adults were better at detecting 9-year-old and adult witness's truthfulness but not for the 4-year-old (Newcombe & Bransgrove, 2007). These findings suggest that other factors may take priority in adults' decision-making about children's credibility. The following sections describe these other factors that may influence adults' perceptions.

Confidence. Several studies have reported witness perceived confidence as one of the strongest predictors of perceived accuracy and credibility ratings (Brewer & Burke, 2002). Adult mock jurors consistently use children's displayed confidence as an indicator of their credibility. For example, Fawcett and Winstanley (2018) found confident 8- to 16-year-old alibi witnesses were perceived as more honest, accurate, and reliable than unconfident alibi witnesses. Connolly et al. (2008) found children who experienced a repeated event were less confident than those who experience a single episode of the event. As a result, those children with repeated exposure were seen as less credible. Similar findings have been found for adult witnesses. In a study that investigated the relationship between perceived confidence and accuracy in adults' cued recall of a video-recorded event, participants were perceived as significantly more confident when recalling accurately than when recalling inaccurately (Vredeveldt & Sauer, 2015). However, despite findings showing adults' perceptions of witness confidence to relate to credibility, the link between actual accuracy and perceived confidence is weak (Shaw & McClure, 1996).

Consistency and accuracy. In addition to perceived confidence, adults also use the consistency and perceived accuracy of children's testimony as a gauge to determine their credibility. For example, 4- to 7-year-olds' reports of details of a repeated event were judged as less consistent than details of a unique event; and in turn, their reports were judged as less

credible (Connolly et al., 2008, Exp 1). When the number of descriptive errors was manipulated, Bruer and Pozzulo (2014) found witnesses aged 4, 12, and 20 who made no errors in their description of a mock crime event and perpetrator were judged by adult participants to have higher overall scores in reliability and accuracy than those who made six errors. In a study that examined sexual abuse reports by 3- to 16-year-olds, consistency was predicted by their memory abilities (Ghetti, Goodman, Eisen, Qin, & Davis, 2002). It was also found that older children's reports were more consistent than their younger peers. These findings demonstrate that the consistency of reports may be an indication of developmental ability that increases with age. Together, these findings indicate that consistency and accuracy, which often increase with age as well as cognitive capabilities, may play a significant role in jurors' decision-making of witness credibility.

The present study

In this study, we examined whether adults' perceptions of children's credibility in recalling temporal details of an event are influenced by whether a timeline was used. Four interviews from children aged between 7 and 9 were used, each with a different child responding to temporal questions regarding their experience at an on-going summer camp. Adult participants watched one of the four interviews, with each interview separated into two halves, half with the timeline (timeline) and the other half without the timeline (verbal). Both half-interviews centered around the same questions related to target details from their camp experience and differed only by interview style (timeline vs. verbal). After watching the first half of the interview, participants rated their perceptions of the child's cognitive competence, interview performance, confidence, honesty, effort, suggestibility, cooperativeness, and overall credibility. Another questionnaire followed the second half of the interview (video 2) and asked for adults' perceptions of the same

factors in comparison to the first half of the interview. Participants either watched the timeline interview first or the verbal interview first, resulting in two interview conditions, namely *verbal first* and *timeline first*. Also, after the timeline interview video was watched, participants provided ratings on children's understanding, helpfulness, and overall impact of the timeline.

This study is novel in several ways. First, no published work, to our knowledge, has examined adults' perceptions of any visual aid used in children's recall of a repeated event. Second, as several studies have cautioned that the use of any props should follow exhaustive verbal recall (e.g., Strange et al., 2003; Salmon et al., 2012), the current study included the order of interviews as the condition for assessing whether the timing of the introduction of the timeline affected adults' perceptions. Our hypotheses for adults' perceptions of children's interview characteristics and overall credibility are as follows:

1) Adults in the current study would have different perceptions of cognitive competence, interview performance, other interview characteristics, and overall credibility for children who used the timeline compared to those who did not. This follows from previous rationales for introducing anatomical dolls and human figure diagrams, which were originally introduced, because it was believed that such props could aid children's interview experience.
2) A change in adults' perceptions after watching the second half of the interview regardless of interview conditions because adults were expected to have different perceptions for the timeline and verbal interviews.

Furthermore, we explored whether adults' perceptions of the timeline as an interview aid would be related to how they would view children's credibility. The main objectives of the current study were to explore whether adults' perceptions of children's credibility differed by the

presence of the timeline and whether those perceptions are in line with empirical research showing the potential risks that timelines, along with other interview aids have on children's recall.

Method

Participants. Two hundred and twenty undergraduate students were recruited through the university psychology experience program. In total, data from 45 participants were excluded from the study; 23 were excluded because their answers to the same two questions regarding the child's accuracy were conflicting (i.e., rated "slightly accurate" in one question and "slightly inaccurate" in a later question); and 22 did not pass the age manipulation check (i.e., they did not know the age of the child in the video). The final sample included in all analyses comprised of 175 participants (73% female) aged between 17 and 25 (Mean age = 18.94 years, *SD* = 1.36) who attended the same university. Approximately 71% of students were enrolled in a Science major, 13% in Health and Sciences, and the remainder (16%) were enrolled in other majors. Participants primarily identified as White (64%); 16% identified as Asians, 6% identified as Hispanic, and the remaining 14% identified as Black, Indigenous, or Other. More than 90% of the participants reported having a previous interactive experience with children aged between 7- and 9-years-old. The study was reviewed and approved by the university-level research ethics board. Each participant received 0.75 research credit.

Materials. The videos of four children interviewed individually about the target event were used in the present study. Permission to use these videos was granted by their parents from a previous study. Due to a limited number of available videos, these four were selected based on the accuracy and consistency of their reports (see Child interview video section for more details).

Target event. Four children aged between 7 and 9 (Mean age = 7.98 years) were interviewed individually about their experience at the BrainWorx summer camp. The week-long camp began on a Monday; children were interviewed on Thursday of the same week. Each day of the camp was similar in structure and contained 6 target items that were categorized as variable, fixed, or unique. Variable items refer to activities that happened every day but varied in instantiations. For example, children watched a movie (an item) each day of the camp, but each movie they watched was different (instantiations). Fixed items refer to details that were presented in the same way every day. For example, children went outside to play every day of the week. Unique items refer to details that only happened one time. For example, children went rock climbing on the Wednesday only. The remaining five questions were suggestive and referred to details that never occurred. For example, children were asked about the time they played the Scrabble board game at BrainWorx (they never played Scrabble at BrainWorx). See Table 16 for the full list of items and their respective instantiations (where applicable).

Child interview videos. Each child was interviewed first verbally and then with a timeline (verbal-first condition), or vice versa (timeline-first condition). Thus, each interview video was divided into two halves: a verbal and a timeline interview. Eight orders (sets) of presentation were created from the videos with 2 orders per interview (e.g., verbal-first and timeline-first). All videos included the child's age displayed at the beginning of the videos.

Sixteen questions about the target event were administered in random orders for each interview. Eight questions referred to variable items; two questions referred to unique items; one question referred to a fixed item; the remaining five questions were suggestive questions that referred to details that never happened and were asked with misleading information. All questions asked for the day on which an activity or detail of the camp occurred. For example, for

a variable item – 'movie', children were asked "*I know you got to watch MOVIES at BrainWorx. On which day or days of the week did you watch The LORAX at BrainWorx?*". See Appendix B for the full list of interview questions.

The same set of questions was administered during the verbal-only and timeline interviews, which were presented in random orders. Prior to the start of the timeline interview, children were shown the timeline and instructed that the days shown on the timeline represent the days of the week. Children were then asked to point to the first time they were at BrainWorx. All children correctly pointed to Monday. See Appendix C for the timeline used in this study.

Each child's accuracy and consistency were calculated based on the 16 questions asked about the target event. Accuracy was calculated by dividing the number of correct responses by 16; consistency was calculated by dividing the number of consistent responses by 16. Two child videos were high in consistency (M = 0.97) and accuracy (M = 0.77); the other two were low in consistency (M = 0.50) and accuracy (M = 0.50). See Table 17 for the accuracy and consistency for each child video.

Credibility questionnaire. Refer to Appendix D for the full questionnaire for both conditions. Undergraduate students rated, on Likert scales, their perceptions of the child's interview characteristics; namely cognitive competence, interview performance, confidence, honesty, effort, suggestibility, cooperativeness, and overall credibility (Connolly et al., 2008), after watching the first half of interview (Refer to the Video 1 section of the questionnaire in Appendix D). After watching the second half of the interview, participants provided ratings for the same measures in comparison to the first half of the interview (e.g., *"How intelligent do you think the child was, compared to what you saw in the first [half of the] interview?"*). Refer to the Video 2 section of the questionnaire in Appendix D. Additional questions regarding the timeline

were asked after participants had completed the timeline half of the video (Refer to the Timeline section in Appendix D). In other words, in the timeline first condition, participants filled out the timeline section after watching the first video; and in the verbal first condition, participants completed the Timeline section after watching the second video. See Table 18 for the main measures and their respective question number(s) for before-comparison and after-comparison phase.

Design and procedure. Undergraduate students participated in this study in a group of up to 25 individuals. All participants were instructed to bring their laptop or smartphone to complete the in-lab online questionnaire. The questionnaire was developed using Qualtrics ®. Participants first provided demographic information, and then were briefed about the background of the target event (i.e., BrainWorx began on a Monday and the interview was conducted on the Thursday of the same week). They were then instructed that the interviewer was blind to the answers of the interview questions, thus her interactions with the child were simply for verification or to encourage the child to participate in the questioning process. See Appendix E for the instruction script.

A research assistant played the first half of the interview on a projector screen and then instructed participants to complete the first part of the questionnaire (before-comparison phase). Participants were asked to wait for the research assistant's instructions before proceeding to the second half of the questionnaire. The second half of the interview was then played. After the video was played, participants were instructed to complete the second half of the questionnaire (after-comparison phase). Each group of participants was randomly assigned to one child's interview (e.g., Child ID #01) and the condition (verbal first or timeline first). Depending on the order of presentation, participants watched the verbal condition first and then the timeline

condition, or vice versa. Questions about the timeline were only asked in the timeline interview. See Figure 4 for the study procedure.

Coding. Ratings from the second part of the questionnaire were made in comparison to the first part of the interview. Seventeen questions about perceptions of cognitive competence, interview performance, confidence, honesty, effort, suggestibility, cooperativeness, and overall credibility were asked after the second half of the interview video was played. For each question, participants used a 5-point scale to rate how the child's performance compares to the first half of the interview. For example, participants rated 1 for a lot less intelligent, 2 for slightly less intelligent, 3 for about the same, 4 slightly more intelligent, and 5 a lot more intelligent for the question, "How intelligent do you think the child was, compared to what you saw in the first [half of the] interview?". These ratings were recoded into -2, -1, 0, 1, and 2 for a lot less …, slightly less …, about the same, slightly more …, and a lot more … respectively. See Appendix D for the full questionnaire used in each condition.

Results

Preliminary analysis

Preliminary analyses were conducted to eliminate the effect of extraneous variables on the main dependent measures. First, no effect of participant gender emerged on any of the main dependent measures across levels of the independent variable, Fs ≤ 2.74, ps $\geq .10$. Second, the effect of participants' experience with young children (Child Experience variable) was significant for both the before-comparison and after-comparison measures, Fs ≥ 1.604, ps $\leq .02$ Wilk's $\lambda \geq 0.65$. As a result, this variable was included as a covariate in the main analyses where Multivariate Analyses of Covariance (MANCOVAs) were first run to determine if the covariate had significantly affected the dependent measures. Overall results from MANCOVAs were

highly consistent with that of MANOVAs without the covariate. As such, only results from MANOVA are presented.

Descriptives. The scores of before-comparison variables ranged from 1 to 4 for interview performance and suggestibility, with higher values indicating a higher perceived accuracy and coherence (interview performance) and higher resistance against suggestive questions respectively. For all other main measures, ratings ranged from 1 to 5, with higher values indicating that the participants considered the child to be more intelligent (cognitive competence), confident, honest, effortful toward answering questions, cooperative, and overall credibility. After-comparison variables ranged from -2 to 2, with negative ratings indicating poorer perceived cognitive competence, interview performance, confidence, honesty, effort, suggestibility, cooperativeness, and overall credibility for the second interview video in comparison to the first interview video. On the other hand, positive after-comparison ratings indicated better perceptions of the main measures compared to the first interview video. Zero ratings indicated a no-change in perceptions compared to that of the first interview. Perceptions of the timeline included three measures: timeline helpfulness, child's understanding of the timeline, and overall impact of the timeline. Timeline helpfulness and child's understanding of the timeline ratings ranged from 1 to 5, with higher values indicating a more positive perception toward the timeline; and overall impact of the timeline was measured on a -2 to 2 scale, with negative values indicating a negative impact and positive values indicating a positive impact, and zero indicating no impact. See Table 19 for descriptive statistics for the main measures for each interview phase and condition.

The main analyses are presented below by interview phase. To test whether participants' perceptions of the child's interview characteristics and credibility differed by condition (H1), a

direct test of condition differences on participants' perceptions of the child's credibility and interview characteristics was made possible by conducting between-subject analyses on the before-comparison ratings (verbal video vs. timeline video). To test whether the order of the presentation of the interview videos impacted participants' comparison of the two interview videos, their ratings in the after-comparison phase were examined by condition.

Main analyses

The dependent measures analyzed in the main analyses included cognitive competence, interview performance, confidence, honesty, effort, suggestibility, cooperativeness, and overall credibility rated in both the before-comparison phase and the after-comparison phase. See Table 19 for descriptive statistics.

Before-comparison ratings. The analyses of before-comparison ratings tested whether adults' perceptions of the child's interview characteristics and overall credibility differed by whether the timeline was used (H1). All of participants' credibility and interview characteristics ratings of the first interview video they watched were significantly correlated, $r \geq .34$, $p < .001$. See Table 20 for the Pearson correlations. Therefore, the before-comparison dependent measures were analyzed using two one-way Multivariate Analysis of Variance (MANOVA) with condition (verbal-first, timeline-first) as the between-subject variable. The first MANOVA examined interview performance and suggestibility as the dependent variables because they were measured on 4-point scales. No significant effect of condition was found, $F(2, 148) = 0.13$, $p = .88$, Wilk's $\lambda = 1.00$, $\eta_{p2} = 0.002$. Another MANOVA examined cognitive competence, confidence, honesty, effort, cooperativeness, and overall credibility. No significant effect of condition was found, $F(6, 167) = 1.13$, $p = .35$, Wilk's $\lambda = 0.96$, $\eta_{p2} = 0.04$. Participants who watched the verbal video first and those who watched the timeline video first provided comparable ratings on measures of

cognitive competence, interview performance, confidence, honesty, effort, suggestibility, cooperativeness, and overall credibility. See Table 19 for descriptive statistics.

After-comparison ratings. The analyses of after-comparison ratings tested whether adults' perceptions of the child's interview characteristics and overall credibility changed as a result of seeing the two interviews in different orders (H2). Similar to the before-comparison analyses, all of participants' credibility judgments and perceptions of interview characteristics of the second video they watched in comparison to the first video were significantly correlated, $r \geq .245$, $p \leq .001$. See Table 21 for the Pearson correlations. Therefore, the after-comparison dependent measures were analyzed using a one-way MANCOVA with condition as the between-subject variable. No significant multivariate effect was found, $F(8, 165) = .1.58$, $p = .13$, Wilk's $\lambda = 0.93$, $\eta_{p2} = 0.07$. Participants who compared the timeline video against the verbal video and those who compared the verbal video to the timeline video rated comparably on measures of cognitive competence, interview performance, confidence, honesty, effort, suggestibility, cooperativeness, and overall credibility. See Table 19 for descriptive statistics.

Timeline perceptions

A correlation analysis showed that the timeline measures were all significantly correlated, $r \geq .24$, $p \leq .001$. See Table 22 for the Pearson correlations. Therefore, a one-way MANOVA was conducted to test for any effect of condition on participants' perceptions of the timeline. The multivariate effect of condition was not significant, $F(3, 171) = 2.34$, $p = .08$, Wilk's $\lambda = 0.96$, $\eta_{p2} = 0.039$. However, due to the exploratory nature of the study, follow-up independent samples *t*-tests were conducted for each timeline measure with condition as the between-subject factor. Results show a significant effect of condition on timeline helpfulness, $t(173) = -2.51$, $p = .01$, Cohen's $d = 0.38$. Participants who watched the timeline video after the verbal video (the verbal-

first condition) perceived the timeline as less helpful to the child's recall (M = 2.65, SD= 0.12) than those in the timeline-first condition (M = 3.15, SD = 0.11). No other significant effects were found, ts ≤ |1.64|, p ≥ .10.

Exploratory analyses

Due to the study design, the timeline video was shown to participants either before the verbal interview (timeline-first) or after the verbal interview (verbal-first). Specifically, timeline ratings were provided in the before-comparison phase in the timeline-first condition, whereas in the verbal-first condition, the timeline ratings were provided in the after-comparison phase. As such, to examine whether participants' perceptions of the timeline differed based on the order of their watching the interview videos, only the after-comparison credibility ratings can be examined as the outcome variable of interest. To explore the relationship between participants' perceptions of the timeline and their after-comparison overall credibility between the interview videos, a series of moderation analyses was conducted. In each moderation analysis, a timeline measure (timeline helpfulness, child's understanding of the timeline, or overall impact of the timeline) was examined as a potential moderator for the relationship between condition and the after-comparison overall credibility. Table 23 shows the impact of high and low groups of each timeline perception rating on the after-comparison credibility rating. Moderation analyses were conducted using PROCESS v3.3 (Hayes, 2018) with the after-comparison overall credibility as the dependent variable and are presented by each potential moderator as follows.

Timeline helpfulness. The overall model was significant, $F(3, 171) = 9.34$, $p < .001$, $R_2 = .14$. No significant main effect of condition nor timeline helpfulness was found, ts ≤ |1.67|, p ≤ .10, but the interaction between timeline helpfulness and condition was significant, b = -0.60, $t(171) = -5.11$, $p < .001$. The addition of the interaction explained an additional 13% of the total

variance, $F(1, 171) = 26.17$, $p < .001$, $\Delta R^2 = .13$. Among participants who gave a relatively low timeline helpfulness rating on a 5-point scale (1 SD or more below the mean, $\text{helpfulness}_{\text{low}} = 1.78$), the after-comparison credibility rating when the verbal video was watched before the timeline video (verbal-first) was lowered. When the timeline video was watched before the verbal video (timeline-first), participants increased the child's after-comparison credibility rating, $b = 0.60$, $t(171) = 3.26$, $p = .001$. Conversely, for participants who gave a relatively high rating in timeline helpfulness (1 SD or more above the mean, $\text{helpfulness}_{\text{high}} = 4.06$), they increased the child's credibility rating after comparison when the verbal video was watched before the timeline video (verbal-first); and when the timeline video was watched before the verbal video (timeline-first), participants lowered the child's credibility rating after comparison, $b = -0.77$, $t(171) = -3.99$, $p < .001$. No significant condition effect was found for those who rated average in *timeline helpfulness*, $b = -0.08$, $t(171) = -0.63$, $p = .53$. See Figure 6 for a graphical depiction of the moderation effect of timeline helpfulness on the relationship between condition and participants' change in credibility rating.

Child's understanding of the timeline. The overall model was significant, $F(4, 170) = 7.66$, $p < .001$, $R^2 = .12$. No significant main effect of condition nor the child's understanding of the timeline was found, $t\text{s} \leq |0.31|$, $p\text{s} \geq .75$. The interaction effect between timeline understanding and condition was significant, $b = -0.72$, $t(171) = -4.74$, $p < .001$. The addition of the interaction explained an additional 12% of the total variance, $F(1, 171) = 22.46$, $p < .001$, $\Delta R^2 = .12$. For participants who rated the child as having relatively low understanding for the timeline on a 5-point scale (1 SD or more below the mean, $\text{understanding}_{\text{low}} = 2.69$), the after-comparison credibility rating when the verbal video was watched before the timeline video (verbal-first) was lowered. When the timeline video was watched before the verbal video

(timeline-first), participants increased the child's after-comparison credibility rating, $b = 0.63$, $t(171) = 3.31$, $p = .001$. Conversely, for those who thought the child had a relatively high understanding of the timeline (1 SD or more above the mean, understanding $_{high} = 4.49$), their after-comparison credibility rating when the verbal video was watched before the timeline video (verbal-first) was increased. When the timeline video was watched before the verbal video (timeline-first), participants lowered the child's after-comparison credibility rating, $b = -0.67$, $t(171) = -3.51$, $p < .001$. There was no significant condition difference for those who rated average in timeline understanding, $b = -0.02$, $t(171) = -0.14$, $p = .89$. See Figure 7 for a graphical depiction of the moderation effect of participants' perception of children's understanding of the timeline on the relationship between condition and their change in credibility rating.

Overall impact of the timeline. The overall model was significant, $F(3, 171) = 12.32$, $p < .001$, $R^2 = .18$. Neither a significant main effect of condition nor an overall impact of the timeline was found, $ts \leq |0.28|$, $p \geq .780$. The interaction effect between overall impact *of the timeline* and *condition* was significant, $b = -0.92$, $t(171) = -6.02$, $p < .001$. The addition of the interaction explained an additional 17% of the total variance, $F(1, 171) = 36.19$, $p < .001$, $\Delta R^2 = .17$. For participants who perceived a more negative overall impact of the timeline (1 SD below the mean or lower, overall impact $_{neg} = -0.35$), the after-comparison credibility rating when the verbal video was watched before the timeline video (verbal-first) was lowered. When the timeline video was watched before the verbal video (timeline-first), participants increased the child's after-comparison credibility rating, $b = 0.736$, $t(171) = 4.06$, $p < .001$. Conversely, for those who rated the *overall impact* of the timeline more positively (1 SD or more above the mean, impact $_{pos} = 1.34$), the after-comparison credibility rating when the verbal video was watched before the timeline video (verbal-first) was increased. When the timeline video was watched

before the verbal video (timeline-first), participants lowered the child's after-comparison credibility rating, $b = -0.81$, $t(171) = -4.45$, $p < .001$. No significant condition difference was found for those who gave a rating close to no impact, $b = -0.04$, $t(171) = -0.28$, $p = .78$. See Figure 8 for a graphical depiction of the moderation effect of participants' perception of the overall impact of the timeline on the relationship between condition and change in credibility rating.

Discussion

No published work has investigated whether the presence of any visual aid changes adults' perceptions of the credibility of child testimonies. The current study directly examined adults' perceptions of the timeline as an interview aid for 7- to 9-year-old children's recall of temporal details by comparing a verbal interview video to a timeline interview video, as well as two orders in which the interview videos were viewed by the participants. This design allowed, first, a direct comparison of adults' perceptions of a children's recall when interviewed with or without a timeline video (before-comparison analyses). Findings suggest that the timeline did not change adults' perceptions of children's cognitive competence, interview performance, confidence, honesty, effort, suggestibility, cooperativeness, and overall credibility. Second, the condition factor (order of interview) allowed us to examine whether the timing of the introduction of the timeline changed participants' perceptions of children's credibility and recall characteristics. The order of the introduction of the timeline (before or after a no-timeline interview video) did not directly contribute to differences in participants' perceptions of children's recall. However, participants perceived the timeline as more helpful when they had not previously seen the child answering questions verbally (timeline-first condition). Moreover, participants' perceptions of the helpfulness, the child's understanding, and the overall impact of

the timeline moderated the effect of interview order on participants' changes in credibility ratings. Findings are discussed with respect to factors that influence adults' perceptions of children's credibility, as well as forensic implications.

The timeline as an interview aid

When participants had only seen one interview (verbal or timeline), their perceptions of cognitive competence, interview performance, other witness characteristics, and overall credibility did not differ. The presence of the timeline did not lead to participants judging children's memory reports differently. This finding is in stark contrast to a general belief that physical objects used as interview props may be beneficial in child interviews. Visual aids such as anatomical dolls and human figure diagrams were originally introduced by adults in an attempt to help children recall more information and more accurately (see Poole & Bruck, 2012, for a review on the impact of interview aids on children's reports). As such, we expected adults in the current study to show a tendency towards giving more credit to children who used the timeline as a visual aid. Studies 1 and 2 of this dissertation empirically investigated the effect of using the timeline on children's recall of a repeated event and yielded no significant differences in accuracy nor the amount of provided details compared to a verbal interview. As these findings suggest that an interview using the timeline was no different than a non-suggestive and good-practice interview, the timeline may be a redundant tool in child interviews. Adults' perceptions of the timeline in the current study coincide with findings from Studies 1 and 2. The current study suggests that adults strongly associate credible memory reports with one's cognitive competence, interview performance, effort, confidence, and cooperativeness, rather relying on whether a visual aid was present.

Adults' perceptions of the timeline

Participants did not judge children's credibility differently in a no-timeline interview versus a timeline interview when each group had only seen one interview video. However, differences in their credibility judgments emerged after making a comparison between the two interview videos when their perceptions of the timeline were considered. First, participants believed that the 7- to 9-year-old children in the current study had a high understanding of what the timeline represented (mean = 3.59 on a 5-point scale). Second, when the timeline video was presented before the verbal interview video, adults perceived it as more helpful than when the timeline was used after the verbal interview. This is an interesting and important finding showing that adults valued children's verbal-only interview performance and their ability to answer temporal questions without the help of a prop. To look at this from another perspective, it could be that participants initially perceived children as not capable of answering the memory questions and viewed the timeline as helpful. As also demonstrated by the difference in timeline perceptions between the two interview orders; perceptions of the timeline in the timeline-first condition were significantly more positive than in the verbal-first condition. The verbal interview video that followed the timeline video could have demonstrated that the child was well-equipped to answer the interview questions without the help of any visual aid; thus providing lower ratings on perceptions of the timeline's helpfulness and the overall impact on the child's recall compared to when the timeline video was presented first. This finding goes further to support that under a good-practice verbal interview, children could demonstrate the ability to provide a good memory report without the help of any visual aid (e.g., Lamb et al., 2008). Adults holding varying levels of perceptions of a child's credibility could lead to serious consequences in legal settings where children are sometimes interviewed repeatedly and thus evaluated repeatedly for their suitability for the court (e.g., Brown, Pipe, Lewis, Lamb, and Orbach, 2012). As such, it is paramount that

adult decision-makers are well-informed of the effects of visual aids on children's memory reports.

Furthermore, participants' perceptions of the timeline consistently moderated the relationship between the condition (interview order) and their after-comparison credibility rating. Specifically, when the timeline video followed the verbal video (verbal-first condition), participants who viewed the timeline as more helpful and having a positive impact on the child's recall perceived a positive change in credibility between the two interview videos; those who viewed the timeline negatively perceived a decline in the credibility from the verbal video to the timeline video. In contrast, when the verbal video followed the timeline video (timeline-first condition), those who perceived the timeline as more helpful and having a positive impact on the child's recall thought the interview in the verbal video as having a lower credibility rating than in the timeline video; and those who viewed the timeline more negatively gave a positive change in credibility rating from the timeline video to the verbal video. Two important messages are derived from these findings. First, participants' individual differences in their perceptions of the timeline (helpfulness and overall impact) were related to how they judged the child's credibility. Second, although the child's understanding of the timeline also moderated the relationship, this rating was highly associated with the perceived cognitive competence of the child (r = .62). As such, this relationship could suggest that it was the adults' perceptions of the child's cognitive competence that was driving the overall perceived credibility rather than participants' individual differences in timeline perceptions.

Helpfulness and understanding ratings were not consistently related to any measures of cognitive competence, interview performance, nor other witness characteristics. What led to the differences in timeline perceptions remains unknown. Perhaps a study that qualitatively

investigates adults' perceptions of the timeline and other visual aids, in general, could shed light on whether and how these aids could be well incorporated into investigative interview settings. Furthermore, depending on how the participants perceived the timeline, the change in their credibility ratings between the two interview videos was significant. In an investigative setting, it could be risky to rely on the adults' individual differences in their perceptions of the timeline rather than scientific indicators to determine a credible eyewitness report.

Adults' perceptions of credibility

Overall, consistent with eyewitness credibility literature, adults' perceptions of a child's cognitive competence, honesty, effort, cooperativeness, confidence, and resistance to suggestibility were found to be positively associated with their overall perception of credibility. Specifically, higher cognitive competence ratings were associated with higher credibility ratings in the before-comparison phase under both the verbal-first and the timeline-first conditions. In previous literature, higher perceived cognitive competence has been found to relate to higher credibility. For example, Connolly et al. (2008) found 4- to 7-year-old children who reported details from an event that occurred only once were seen as more cognitively competent and thus were rated as more credible than those who report an episode from a recurring experience. Similarly, the perceived interview performance (accuracy and coherence) was also related to perceived credibility (Connolly et al., 2008). Accuracy rating is sometimes included as an inherent characteristic of one's cognitive competence (Connolly et al., 2008), but the current study made a distinction between interview performance and cognitive competence because accuracy, in particular, has been studied independently in child eyewitness studies (e.g., Bruer & Pozzulo, 2014). Although adults' ability to discriminate between low- and high-accuracy testimonies is weak (e.g., Connolly et al., 2008; Leippe et al., 1993), they often use their

perceived accuracy as a gauge to determine the overall credibility. With respect to adults' perceptions of children's honesty in the study, the presence of the timeline did not affect the honesty rating; all children in the interview videos were rated as highly honest (3.74 on a 5-point scale). Although the current study asked children to recall a fun and non-stressful event, and thus honesty is not particularly salient because the motivation to lie is very low, children's honesty was also found to be positively associated with perceived credibility. The current study replicated the results of Connolly et al. (2008), which also focused on the perceived credibility of a repeated-event recall, that consistently showed a positive relationship between cognitive competence (including accuracy) and overall credibility, as well as between honesty and overall credibility. The current study provides further evidence to support that perceived cognitive competence, accuracy, and honesty are strong indicators of credibility, and that adults tend to judge the credibility of memory reports of a repeated event based on these perceptions, rather than solely based on the presence of the interview aid – the timeline.

In line with previous research, other witness characteristics were also strongly associated with perceived credibility. Specifically, confidence, effort, and cooperativeness were all positively associated with perceived credibility. Although the correlation between confidence and actual accuracy is weak (e.g., Goodman et al., 2002; Newcombe & Bransgrove, 2007; Shaw & McClure, 1996), confidence in reporting has been consistently shown to relate to perceived accuracy and thus higher perceived credibility (e.g., Fisher & Mitchell, 2009). Previous literature has shown mixed results regarding witness effort. Some studies have shown those who are perceived as less effortful are also more honest because they are simply recalling the truth (Evans, Michael, Meissner, & Brandon, 2013; Weinsheimer, Coburn, Chong, MacLean, & Connolly, 2017). However, as mentioned, honesty was not a salient factor in the current study,

and thus a child putting in more effort may be seen as more likeable (Weinsheimer et al., 2017) and having a more serious attitude toward answering interview questions. Moreover, several studies have also shown credibility to be associated with cooperativeness (Fisher & Mitchell, 2009; Weinsheimer et al., 2017). In forensic contexts, cooperativeness may be an indication of volunteering truthful information (Fisher & Mitchell, 2009).

Limitations and implications

The current study helped us understand that individual differences exist in the perceptions of the timeline as an aid and how those differences change adults' perceptions of children's report credibility. Since Study 1 and Study 2 consistently showed that the timeline may not be any more beneficial than a good-practice verbal interview, future research should further examine whether the inclusion of these scientific findings would help adults make more sensible credibility judgments. Specifically, findings such as that of Study 1 and and Study 2 showing limited advantages of the timeline could serve as a guide for decision-makers when making comparisons between interviews that used a visual aid and those that did not.

Like many other visual aids, a timeline can take on many different forms (e.g., horizontal vs. vertical, pictorial cues vs. numbered cues). It is not fair to generalize findings from the current study to scenarios that involve other forms of timelines. However, by including a control (no-timeline) condition, the current study at least demonstrates some evidence that a no-timeline condition is perceived differently (sometimes positively, and sometimes negatively) depending on the timing of the introduction of the timeline. As investigations of child sexual abuse often involve repeated interviewing (e.g., Brown et al., 2012 ; Goodman & Quas, 2008), the inclusion of any visual aid at any point during the investigation could lead to false impressions and result in a misjudgment of the child's credibility. This suggests that the inclusion of a timeline in any

phase of a good quality forensic interview poses a risk of misjudging the credibility of eyewitness reports. In summary, in line with findings from Study 1 and 2, the current study suggests that a verbal interview that follows a good practice is sufficient in extracting forensically relevant information from young children.

Chapter 5

General Discussion

A timeline is a visual representation of time and is often used in educational settings to help children develop concepts about time and continuity, as well as an understanding of the past and future (Hoodless, 2002). Timelines can take on many forms, such as representing time points on horizontal or vertical lines, with or without pictorial cues. While these forms of the timeline are often found in educational settings as instructional tools, they are sometimes also popular in investigative settings as interview aids. Investigative interviewers have always been in search of special techniques that would help them elicit forensically relevant information from child witnesses (Poole & Bruck, 2012). One of those techniques is using visual aids. Specifically, timelines could be used for retrieving information about specific temporal episodes.

The three studies in this dissertation collectively sought to address the central question of whether the timeline is an appropriate visual aid for 6- to 10-year-old children. Visual aids such as anatomical dolls and human figure drawings have been shown to pose risks to the quantity and quality of children's recall. The timeline, however, has been used in investigative settings but rarely been studied using child samples. This dissertation comprises a systematic evaluation of the role of the timeline in two memory-retrieval settings. First, by examining the effect of the timeline on children's recall of a repeated event in an interview that included only invitational-prompts (i.e., "*Tell me everything about …*") and second an interview that only asked focused-recall questions; third, by examining adults' perceptions of children's recall of a recurring event in an interview setting with and without the use of the timeline.

Study 1 examined children's responses to a series of focused-recall questions that directed the questions to specific aspects of an instance of a repeated event (i.e., "*What was the*

color ... "). Overall, the timeline did not improve the accuracy of children's recall in response to questions regarding a specific instance and, when asked to make a source decision about details from specific instances, the timeline did not help children improve their source accuracy. Several interesting findings emerged, however. First, a suggestibility effect was more pronounced for the recall of fixed items for children who used the timeline than when no timeline was used. Fixed details are details that appear consistently throughout a repeated event (e.g., the leader of the Activities always wore a red cloak). Research has shown that memory of details that always occur in a repeated experience tends to be much stronger than those that vary (Powell et al., 1999). The alarming finding that the timeline led to a higher suggestibility rate suggests that the timeline may not be a suitable memory retrieval aid in cases where false suggestions have taken place. In forensic settings where retrieving accurate testimonies is of high priority, this finding adds to the current pool of literature that recommends that nonverbal props should always be used with caution. Second, children who used the timeline intruded details from an occurrence closer to the correct occurrence; whereas those in the control condition intruded details from an occurrence further away from the correct occurrence. Although it could not be determined whether making an intrusion error from a closer occurrence would be impactful to a child's overall testimony, the timeline was systematically creating a different mental representation of the repeated event than when it was not used.

Study 2 examined children's spontaneous recall of a repeated event (i.e., "*Tell me everything you remember about the Laurier Activities*") using the same repeated-event paradigm as in Study 1. Children's spontaneous recall of target items were coded and analyzed according to Brubacher et al. (2012) by detail level (item-level vs. instantiation-level) as well as by specificity (episodic vs. generic). The purpose of this study was to examine whether the timeline,

displaying all occurrences of the repeated experience, could lead to more reports of specific details from a specific occurrence. Overall, children who used the timeline did not show an increase in the report of instantiation-level nor episodic details. These findings provide additional evidence to support our conclusions from Study 1 where the timeline did not help with children's recall. Additional findings replicate findings from previous work. First, children were responding in a style corresponding to the interviewer's question (i.e., general recall questions elicited more generic responses and target recall questions elicited more responses about the target episode). Second, all children were more likely to recall at an item-level when asked to recall *"everything [they] remember about the Laurier Activities"*; and recall at an instantiation-level when asked to recall *"everything [they] remember about the time [they] got the candy badge at the Laurier Activities"*. Previous research has demonstrated that when 4- to 16-year-olds were asked about a previous experience, their response style matched with that of their interviewers' (Schneider et al., 2011).

Finally, children's age in Study 2 correlated significantly with the number of specific responses they provided via open-ended recall. Specifically, the number of instantiation-level details in both recall phases and the number of episodic details in the target recall phase increased with age. Age has been shown to be a major factor in determining children's abilities to recall specific details from an instance of a repeated event or make source decisions about details in a repeated event in several studies (e.g., Brubacher et al., 2011; Connolly & Price, 2006). Moreover, these age-related increases also correspond to results from Study 1 that demonstrated that older children were more accurate in recalling details from a specific instance and making source decisions. Whether through open-ended recall or focused-recall questions, findings converge to show that child's ability to recall specifically is largely related to their age.

Together, findings from Study 2 replicated those reported in the limited number of previous laboratory studies that examined children's narrative recall of a repeated experience (Brubacher et al., 2012). The nonsignificant differences between the timeline and the control group serve as evidence that responding to open-ended questions using nonverbal props such as the timeline, in this case, did not help children recall more specific instantiations nor details about a specific episode.

Results from Study 1 and 2 prompted the further question of whether adults' understanding of the role of the timeline is in line with that of research findings showing no support for its use. Study 3 investigated adults' perceptions of a group of children using a timeline in comparison to a no-timeline control group. In addition, the timing of the introduction of the timeline (before or after a no-timeline interview) was also investigated.

Adults' perceptions of children's interview characteristics correlated with their perceptions of children's overall credibility, which is in line with existing perceived credibility literature that consistently shows this relationship (e.g., Connolly et al., 2008). With regard to comparisons made between the timeline and the verbal interview videos, the present findings are novel and suggest that this is an area worthy of further research. First, adults' perceptions made based on a timeline video and a verbal video did not differ on the dimensions of cognitive competence, interview performance, and overall credibility. Second, differences in adults' perceptions of the timeline led to changes in the overall credibility rating between the timeline and verbal video. Specifically, if adults viewed the timeline as more helpful, better understood by the child, and having a positive impact on the child's recall, then the interview in the verbal video that followed the timeline video was viewed as less credible; and conversely the interview in the timeline video that followed the verbal video was seen as more credible. Similarly, if

adults viewed the timeline less favorably (i.e., lower ratings in helpfulness and child's understanding, and a negative impact on recall), then the interview in the timeline video that followed the verbal video was seen as less credible; and the interview in the verbal video that followed the timeline video was seen as more credible. In summary, a child's overall credibility rating was determined by adults' perceptions of the timeline, as well as the order in which the interviews were watched.

Overall, results do not support the use of the type of timeline used in the present study in investigative interviewing settings. As with findings from research on other interview aids such as anatomical dolls and human figure drawings, the present timeline did not consistently improve children's recall accuracy nor quality when compared to a control group. In addition, when faced with interviews in which recall accuracy was equivalent, adults held different perceptions of the child's credibility when the timeline was used in recall compared to when it was not used.

The studies in this dissertation utilized a well-established repeated-event paradigm and extended the pool of visual aid research by examining perceptions of the timeline from the undergraduate students' point of view. These findings build upon our existing knowledge of how visual aids may affect children's recall and create a pathway for future research on whether using objects representing different concepts (timelines = time, dolls = body) indeed helps concrete thinkers engage in verbalizing information and provide effective retrieval cues (Russell, 2008). According to Piaget's theory of cognitive development, children between ages 7 and 11 are in the concrete operational stage where they begin to apply logic to physical objects and make sense of rules and logical operations. Children in the current studies who were between ages 6 and 10, may very well fit into this stage of concrete operations where they need physical objects for making sense of the abstract world around them. Implications are discussed in detail in the

following sections with respect to: 1) the timeline as an interview aid, and 2) adults' perceptions of the timeline and children's recall competency.

The timeline as a visual aid

In both Study 1 and 2, children who used the timeline did not outperform those who were in the no timeline group in terms of overall recall accuracy, the ability to attribute details to the correct instance, nor the amount of spontaneous reporting of specific target details from the target event or target instance. In other words, having all instances of the target event displayed visually on a timeline did not improve children's recall in any way. Implications are discussed in relation to their application and related theories. First, as timelines take on many forms, the particular format of the timelines used in the current study may pose constraints on our interpretation of the results. Second, children's representational understanding of objects representing other objects or concepts is required before timelines (and other visual aids) could be used as retrieval tools. However, the 6- to 10-year-old children in the current studies should have already passed the stage where symbolic representation poses any issues. Third, as visual timelines represent time concepts, children's temporal understanding is essential in making temporal decisions and recalling memories about time. Finally, as studies in this dissertation aimed to examine whether the timeline serves as an appropriate retrieval aid for particularization (i.e., recalling details from specific episodes), theories regarding children's development of episodic memory are discussed.

Design of the timeline. The physical timelines used in these studies visually demonstrated all occurrences of the activity days using numbers. The introduction of the timeline was carried out in a non-suggestive manner that explained to children what each component of the timeline represented. Namely, "*the number 1 represents the first time you were at the Laurier*

Activities ...". Bruck et al.'s (2000) anatomical doll study suggested that when specific parts of a doll that represented the corresponding body parts are explained in detail, children suggestively reported body touches that did not occur. However, a similar study by Saywitz et al. (1991) did not find a high rate of false reporting. Poole and Bruck (2012) attributed this difference to the labelling of body parts. Specifically, they argued that the labelling of body parts as part of the introduction of the doll in Bruck et al. (2000) prior to recalling touch-related details using the doll may have led to children giving more false reports because children were primed with the labels of body parts whereas in Saywitz et al. (1991), no such introduction was included. For this reason, as part of creating a non-suggestive interview environment, the timeline was introduced with minimal instructions. In child abuse cases, the timeline could be presented differently based on the uniqueness of each allegation such as the frequency of abuse and the length of delay between occurrences. Nonetheless, as recommended by an evidence-based interview protocol -- the NICHD protocol (Lamb et al., 2007) -- it is always recommended that interviewers only provide information (whether verbally or through the use of any nonverbal aid) that was previously mentioned by the child. In the case of the timeline, the numbering and cues provided to the child should correspond to previously mentioned information. Future research could address this issue by examining other variations of the timeline such as a child-generated timeline based on children's labels (rather than an interviewer-generated label) representing specific occurrences. Past research has shown that 7- to 8-year-old children were highly accurate with attributing details that only occurred once during a repeated experience to the correct occurrence, suggesting that these details could serve as a meaningful label for children to talk about the specific occurrence (Brubacher et al., 2012). Further, translating these child-generated

labels to a visual format that contains only information suggested by the child could minimize the amount of interviewer input and allow the child to recall non-suggestively.

Symbolic representation. Children's ability to use an object as a representation of other objects or concepts, known as symbolic representation, has been examined mostly with children under 5. For example, 2 ½ to 4-year-olds have difficulty grasping the self-doll relation (doll's body = self) (Deloache & Marzolf, 1995). Three to 5-year-olds cannot identify the location of objects in a space using a map (Uttal, 2000). The development of full pictorial understanding takes several years (Deloache, 1996; Deloache et al., 2003). In addition, as preschoolers between 3- to 6-years-old go through profound neural and cognitive development (Newcombe et al., 2007), we would assume that the 6- to 10-year-old children in the current program of research should have a relatively good grasp of the symbolic nature of the timeline. However, some interesting findings indicated that such an understanding of symbolic representation may have had effects on children's recall and mental representation of the event. Specifically, results from Study 1 showed that the timeline seemed to have "pushed" the separate instances of the repeated event closer together in children's mental representation of the entire experience (thus leading to potentially more source confusions in the future). This calls for research to examine whether the full potential for the timeline to be used as a visual aid develops in stages. Namely, the stages could involve first the development for an understanding of basic temporal concepts, then an understanding of symbolic representation (timeline = time), and finally the ability to incorporate both understandings into utilizing the timeline as a memory retrieval aid.

Temporal understanding. Research in the past has extensively examined children's temporal understanding through various forms of timelines (e.g., Friedman, 1993; Zhang & Hudson 2019). Although these were not directly testing the timeline as a memory aid, findings

indicated that such a graphical representation of time could be incorporated into children's demonstration of knowledge for time. For example, 5-year-olds could order a series of events using pictorial cues along a timeline of past and present (Busby, Grant, & Suddendorf, 2009). The ability to demonstrate knowledge for "yesterday" and "last week" and their relative ordering on a timeline emerges around age 4 (Tillman et al., 2017). Despite this evidence suggesting that the representational nature of timelines is somewhat understood at an early age, young children's ability to incorporate that understanding into their recall of a past event requires further investigation. The current program of research taps into that relationship by systematically examining the timeline against a no-timeline memory retrieval setting and found no benefit of the timeline on children's recall. Future research could further explore the role of the timeline in recall settings by first establishing children's representational understanding of the timeline and associate that with their memory performance using the timeline.

Episodic memories. The intent of including a timeline in child memory interviews is to facilitate children's episodic recall. In child sexual abuse investigations, a greater number of episodic reports could significantly substantiate child sexual abuse allegations and make children appear more credible (Schneider et al., 2008). Study 1 and Study 2 of this dissertation demonstrated that the accuracy and quantity of recall of a specific occurrence were not affected by the use of the timeline. This nonsignificant finding could be explained by the encoding specificity principle (Tulving & Thomson, 1973) that suggests the successful retrieval of episodic memory depends on whether the retrieval cue "matches" that of the condition under which the memory trace was encoded. The visual depiction of different occurrences on the timeline in this study may not match children's mental representation of those occurrences. As such, the timeline may not have acted as a matching cue for children to retrieve information from

the specific episode. Research on other visual aids such as human figure drawing also tends to find no recall difference between visual aid and control condition (Brown et al., 2007). It could be that the symbolic representation of the figure drawings did not match entirely as the actual touches received by children, which in turn led to no additional advantage for the recall of touches. Further evidence in Study 2 confirmed that the style of questioning (episodic vs. generic) was directly related to how children responded (e.g., Brubacher et al., 2012; Schneider et al., 2015), and was not related to whether the timeline was used. While this is evidence suggesting that the use of episodic interview language may help children recall more episodically, forensic interview questions should remain open-ended (i.e., invitational prompt) and non-misleading to encourage accuracy.

Some expected developmental trends also emerged. Older children in the current studies generally outperformed their younger peers in episodic tasks regardless of the interview condition (control vs. timeline). This age trend was expected based on past research showing children's ability to recall episodically increases dramatically, along with increases in other cognitive functions such as executive function (e.g., Brubacher et al., 2012; Earhart & Roberts, 2014). The ability to recall episodically is also more pronounced in the event of a recurring experience compared to a single experience (Brubacher et al., 2012). This age trend is also supported by Fuzzy-Trace theory which suggests that younger children's verbatim traces (specific and episodic details that contain more than just the gist) decay more rapidly than the traces of older children (Brainerd & Reyna, 1998, 2004). However, the timeline in Study 1 and Study 2 did not seem to enhance children ability to retrieve these verbatim memory traces. Neither younger or older children in the current studies benefited from the presence of the timeline in the retrieval of specific and episodic details. The source-monitoring framework also

helps explain why there was a clear age trend in the report of specific details. The development of source monitoring is gradual across childhood with substantial improvements between ages 3 and 8 (see Roberts, 2002, for a review). Source monitoring involves engaging in higher-order decision-making processes to make source decisions (e.g., was the puzzle a snowman on the first day or the second day?). Earhart (2017, Study 1) provided evidence that as older children demonstrated higher working memory capability, their source monitoring ability also increased for both easy and difficult tasks.

Going back to the intended function of the timeline, which is to help increase the number of episodic reports from younger children who may have more difficulties verbalizing or recalling a past experience than their older peers, the timeline clearly did not fulfill such a function. Children as young as 6-years-old in Study 2 provided episodic reports regardless of whether the timeline was used. Together there is evidence that age and cognitive competence are factors that contribute to children's ability to recall episodically, while the inclusion of the timeline aid played no role, suggesting there may not be a need for such a tool.

Adults' perceptions of the timeline and children's recall competence

Study 3 explored two distinct questions: 1) do adults hold different perceptions toward child interviews that involved the timeline compared to those that did not? 2) And do those perceptions change when they compare a timeline interview to a verbal interview? Results showed that when adults had no comparison, they held similar perceptions of the child's cognitive competence, interview performance, confidence, honesty, effort, cooperativeness, and overall credibility between the two interview techniques. However, once they had seen an interview video of a different technique (with the timeline or without the timeline), their perceptions of credibility changed dramatically based on their view of the timeline. Specifically,

when adults had seen a child demonstrating the ability to answer verbally, they may have believed that the child was well-equipped to answer questions without any external aid. Additionally, when adults viewed the timeline video first, they held a belief that the timeline was much more helpful and believed that the verbal interview that followed was much lower in credibility. This is an important finding that suggests that adults may hold a pre-existing belief that children are not competent in these memory tasks and are in need of some kind of nonverbal aid rather than a belief that they *are* competent. Cleveland and Quas (2016) suggested that adults are sometimes insensitive to children's developmental progression and abilities to provide temporal details. Findings from this study also suggest that children's verbal interview videos could serve as information about their abilities to verbalize responses and help adults make more sensible judgments. Moreover, given that Study 1 and Study 2 consistently demonstrated that the timeline had no or even negative effects on children's recall of a specific instance, it is risky that adults sometimes hold an erroneous belief that the timeline may be helpful. In investigative settings where a timeline is used, such a false or exaggerated belief could undermine a credible testimony. Taken together with results from Study 1 and Study 2 showing that children are indeed capable of answering questions verbally, adults should be informed about children's developmental characteristics and capabilities before asking them to make critical judgements about children's credibility.

Conclusion

Findings from all three studies in this dissertation collectively suggest that the timeline should be used with caution. As with other types of visual aids, the timeline posed a risk of false reports (Study 1) or did not affect the amount of spontaneous recall (Study 2). These measures represent the quality and quantity of an eyewitness report, which should be of the highest priority

in forensic investigations. However, these were not improved by the timeline. The inclusion of a control group in both Study 1 and 2 provided further evidence that children between 6- and 10-years-old are indeed capable of providing quality verbal reports (Lamb et al., 2008).

Furthermore, despite some research suggesting that adults are sometimes insensitive to children's age by discriminating credibility of their report mainly based on perceived confidence (Cleveland & Quas, 2016), adult participants in Study 3, who were between 18 and 25 years of age and would qualify to serve as jurors in Canada, gave due credit to children's demonstration of verbal recall ability. Specifically, it was demonstrated that they valued children's ability to provide verbal-only reports by raising their perceptions of children's credibility after comparing a verbal-only interview video to an interview video where a timeline was used.

Given the findings of the current dissertation, a cautious approach is recommended when investigators use timelines with young children for memory retrieval purposes. These findings 1) emphasize that young children are capable of recalling verbally without the help of any external aid; 2) highlight the importance of using a non-suggestive and quality verbal interview procedure where children are given the chance to exhaust their verbal recall and report everything they remember in a non-suggestive and supportive environment without the use of any nonverbal props; and 3) caution that the development of new effective interview strategies for criminal investigations involving young victims of repeated abuse should be built upon a good exhaustive verbal interview. Any potential incremental benefit of nonverbal props such as the timeline is minimized when a good-practice verbal interview protocol is followed.

Tables

Table 1

Fixed and Variable Items and the Corresponding Suggestion Type in the Biasing Interview for Study 1 and 2

	Target item	Item type	Suggestion set 1	Suggestion set 2
1.	Activity boxes	Fixed	True	False
2.	Leader's cloak	Fixed	False	True
3.	Sticker	Variable	False-internal	False-external
4.	Location of sticker	Variable	False-external	False-internal
5.	Children to sit on X	Variable	False-internal	False-external
6.	Puzzle	Variable	False-external	False-internal
7.	Sheet for puzzle	Fixed	True	False
8.	Sticky cards	Variable	False-internal	False-external
9.	Location of sticky cards	Fixed	False	True
10.	Human body story	Variable	False-external	False-internal
11.	Bookmark	Variable	False-internal	False-external
12.	Refreshment	Fixed	True	False
13.	Free drawing topic	Variable	False-external	False-internal
14.	Utensil	Fixed	False	True
15.	Connect-the-dots	Variable	False-internal	False-external
16.	Container	Variable	False-external	False-internal

Note: *False-internal* suggestions were experienced instantiations from non-target occurrences. *False-external* suggestions were created and never experienced.

Table 2

Entire Set of Target Items and Instantiations for Study 1 and 2

		Instantiations					
Target item		A	B	C	D	E	False suggestion
1.	Activity boxes	Clear	Clear	Clear	Clear	Clear	Purple
2.	Leader's cloak	Red	Red	Red	Red	Red	Yellow
3.	Sticker	Sheep	Piggie	Bull	Horse	Chick	N/A
4.	Sticker	Forehead	Hand	Chin	Cheek	Nose	N/A
5.	Children's seat	Face cloth	Folder	Paper plate	Number mat	Sponge mat	N/A
6.	Puzzle	Scarecrow	Snowman	Castle	Clown	Beach ball	N/A
7.	Sheet for puzzle	Blue sheet	Blue sheet	Blue sheet	Blue sheet	Blue sheet	Black sheet
8.	Sticky cards	Sports	Animals	Transportation	School supplies	Food	N/A
9.	Location of sticky cards	Leader's back	Leader's back	Leader's back	Leader's back	Leader's back	Leader's pants
10.	Human body story	Eating and pooping	Breathing	Bones and muscles	Brain power	Senses	N/A
11.	Bookmark	Hearts	Squares	Circles	Stars	Triangles	N/A
12.	Refreshment	Water mist	Water mist	Water mist	Water mist	Water mist	Hand Sanitizer
13.	Free drawing topic	School	Recess	Family	Self	Friends	N/A
14.	Utensil	Marker	Marker	Marker	Marker	Marker	Crayon
15.	Connect-the-dots	Starfish	Pumpkin	Leaf	Bell	Ice cream	N/A
16.	Container	Pencil case	Ziploc	Glass jar	Envelope	Basket	N/A

Table 3

Means and Standard Deviations for Children's Age in Years by Age Group and Interview Condition for Study 1

	Younger		*Older*	
	Control	*Timeline*	*Control*	*Timeline*
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Age in years	7.34 (0.44) *n* = 31	7.46 (0.42) *n* = 29	8.87 (0.40) *n* = 29	8.94 (0.45) *n* = 32

Table 4

Coding Scheme for Study 1

Code	Definition
Correct response	The instantiation for the target instance was reported.
Internal intrusion error	An instantiation from a non-target occurrence was reported (i.e., the 1st, 2nd, or 3rd) or if children provided more than one instantiation (e.g., *"the puzzle was a castle or clown")* as at least one of the instantiations must have been from a non-target occurrence.
External intrusion error	An instantiation that was not suggested or experienced.
Report of suggestion	The false suggestion given in the biasing interview was inaccurately reported as being in the target instance. For *fixed* items, all suggestions referred to details that children did not experience. False suggestions reported about *variable* items were further classified as referring to the *false-external* or the *false-internal* suggestions.
'I don't know/ remember' response	"I don't know" or "I don't remember" responses
Other	Confusions made for details within an occurrence. *Other* responses were rare and not further analyzed.

Table 5

Means and Standard Deviations for Accuracy by Age Group, Suggestion Type in the Biasing Interview, and Interview Condition for Study 1

	Younger		*Older*	
	Control	*Timeline*	*Control*	*Timeline*
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Fixed items				
True-suggestion	.95 (.15)	.98 (.09)	.97 (.10)	.97 (.10)
False-suggestion	1.00 (.00)	.92 (.15)	.97 (.10)	.86 (.22)
Variable items				
False-internal	.33 (.24)	.36 (.23)	.41 (.17)	.41 (.24)
False-external	.38 (.24)	.32 (.25)	.46 (.29)	.43 (.22)

Table 6

Means and Standard Deviations for The Proportion of Reported Suggestions by Age Group, Suggestion Type in the Biasing Interview, and Interview Condition for Study 1

	Younger		*Older*	
	Control	*Timeline*	*Control*	*Timeline*
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Fixed items				
False-suggestion	.00 (.00)	.02 (.09)	.01 (.06)	.05 (.17)
Variable items				
False-internal	.29 (.21)	.26 (.23)	.23 (.19)	.28 (.24)
False-external	.05 (.09)	.03 (.09)	.01 (.05)	.08 (.10)

Table 7

Means and Standard Deviations for the Proportion of Internal Intrusion Errors for Variable Items by Age Group, Suggestion Type in the Biasing Interview, and Interview Condition for Study 1

	Younger		*Older*	
	Control	*Timeline*	*Control*	*Timeline*
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Variable items				
False-internal	.21 (.15)	.23 (.22)	.18 (.15)	.16 (.11)
False-external	.39 (.22)	.43 (.28)	.31 (.27)	.38 (.23)

Table 8

Means and Standard Deviations the Proportion of Intrusions from Each Previous Occurrence by Age Group, and Interview Condition for Study 1

	Younger		*Older*	
	Control	*Timeline*	*Control*	*Timeline*
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Occurrence 1	.30 (.28)	.13 (.14)	.17 (.27)	.17 (.32)
Occurrence 2	.16 (.22)	.21 (.23)	.48 (.34)	.28 (.30)
Occurrence 3	.54 (.33)	.66 (.27)	.34 (.34)	.55 (.34)

Table 9

Means and Standard Deviations for Accuracy, Average Distance of Errors, and the Proportion of Accurate Attributions for Each Occurrence in Response to the Instance Attribution Questions, by Suggestion Type in the Biasing Interview, Age Group, and Interview Condition for Study 1

	Younger		*Older*	
	Control	*Timeline*	*Control*	*Timeline*
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Accuracy				
False-internal	.30 (.23)	.39 (.18)	.38 (.17)	.38 (.20)
False-external	.36 (.17)	.34 (.19)	.36 (.20)	.33 (.21)
Average distance of erroneous attributions				
False-internal	1.54 (0.47)	1.43 (0.51)	1.48 (0.36)	1.46 (0.43)
False-external	1.59 (0.34)	1.63 (0.38)	1.53 (0.48)	1.58 (0.47)
Overall	1.59 (0.32)	1.54 (0.30)	1.45 (0.41)	1.51 (0.30)
Proportion of accurate attributions				
First occurrence	.35 (.37)	.24 (.32)	.41 (.33)	.33 (.33)
Middle occurrences	.32 (.20)	.40 (.22)	.28 (.19)	.32 (.24)
Last occurrence	.66 (.57)	.78 (.49)	.86 (.57)	.78 (.51)

Table 10

Means and Standard Deviations for Children's Age in Years by Age Group and Interview Condition for Study 2

	Younger		*Older*	
	Control	*Timeline*	*Control*	*Timeline*
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Age in years	7.38 (0.66) *n* = 18	7.40 (0.81) *n* = 18	9.04 (0.42) *n* = 19	9.17 (0.52) *n* = 16

Table 11

Specificity Coding for Study 2

Combination of grammatical tense and temporal marker	Specificity
No tense + No temporal marker	Generic
Could/would/should/got to/had to + No temporal marker	Generic
Can/have/will/get to + No temporal marker	Generic
Present tense + No temporal marker	Generic
Could/would/should/got to/had to + Generic marker	Generic
Can/have/will/get to + Generic marker	Generic
Past tense + Generic marker	Generic
Present tense + Generic marker	Generic
Past tense+ No temporal marker	Episodic
Could/would/should/got to/had to + Episodic marker	Episodic
Can/have/will/get to + Episodic marker	Episodic
Past tense+ Episodic marker	Episodic
Present tense + Episodic marker	Episodic

Table 12

Outliers Removed From the Main Analyses in Study 2

Coding category	Recall phase		Number of outliers	Cutoff value (+/- SD)	Outlier value(s)
Detail level	Target	Number of Instantiation-level details	1	10.04	11
Specificity	Target	Number of Generic details	2	5.59	6
	General	Number of Generic details	2	12.98	13
		Number of Episodic details	1	7.95	9

Table 13

Means and Standard Deviations by Interview Condition and Specificity in Study 2

Recall phase	Interview condition	Episodic details			Generic details		
		N	*Mean*	*SD*	*N*	*Mean*	*SD*
General	Control	36	1.36	1.59	36	4.06	2.32
	Timeline	34	1.94	2.15	33	4.61	2.57
Target	Control	37	3.3	2.25	37	0.59	1.28
	Timeline	34	3.85	2.2	32	1.03	1.31

Table 14

Bivariate Correlations Between Age in Years and Main Dependent Measures in Each Coding Category for Study 2

Coding category	Recall phase		*r*
Detail level	General	Total Item-level	.063
		Total Instantiation-level	**.250***
	Target	Total Item-level	.080
		Total Instantiation-level	**.273***
Specificity	General	Episodic	.061
		Generic	**.276***
	Target	Episodic	**.299***
		Generic	.069

Note: **p* < .05.

Table 15

Standardized Coefficients and Test Statistics for Age in Years and Interview Condition as Predictors in Study 2

	Dependent measure	Predictor	Beta	*t*	*p*
General recall	Item-level	Age in years	.063	.53	.60
		Interview condition	.193	1.63	.11
	Instantiation-level	Age in years	.250	2.13	.04
		Interview condition	.043	0.37	.72
Target recall	Item-level	Age in years	.080	0.68	.50
		Interview condition	.224	1.91	.06
	Instantiation-level	Age in years	.272	2.33	.02
		Interview condition	.102	0.87	.39
General recall	Episodic	Age in years	.059	0.49	.63
		Interview condition	.154	1.28	.21
	Generic	Age in years	.272	2.31	.02
		Interview condition	.104	0.88	.38
Target recall	Episodic	Age in years	.299	2.60	.01
		Interview condition	.125	1.09	.28
	Generic	Age in years	.065	0.54	.59
		Interview condition	.167	1.38	.17

Table 16

Target Items and Respective Instantiations (If Applicable) for the Interview Materials Used in Study 3

Item type	Item	Instantiation
Fixed item	Playing outside	NA
Unique items	Rock climbing	NA
	Red-balloon room	NA
Variable items	Theme of the day	Senses
		Reflexes
		Brain
	Movie	Lorax
		Flubber
		Megamind
	Science experiment	Slime experiment
		Kneejerk experiment

Table 17

Consistency and Accuracy of Interview Videos Used in Study 3

			Consistency	*Accuracy*	
	Gender	*Age in years*		*Verbal video*	*Timeline video*
Child ID 1	Female	7.25	.56	.38	.38
Child ID 2	Female	7.67	.44	.50	.75
Child ID 3	Female	9.25	.94	.75	.69
Child ID 4	Female	7.75	1.00	.81	.81

Table 18

Main Dependent Measures and the Respective Question Number(s) in the Questionnaire for Study 3

Measure	*Question #*
Before-comparison phase	
Cognitive competence	C02, C04
Interview performance	D01, D02
Confidence	D08
Honesty	E01
Effort	D07
Suggestibility	D06
Cooperativeness	D09
Overall credibility	E04
After-comparison phase	
Cognitive competence	F01, F03
Interview performance	F04, F05
Confidence	F11
Honesty	H01
Effort	F10
Suggestibility	F09
Cooperativeness	F12
Overall credibility	H05
Timeline measures	
Child's understanding of the timeline	G01
Helpfulness of the timeline	G02
Overall impact of the timeline	G03, G04

Table 19

Descriptive Statistics for Main Measures by Condition for Study 3

		N	*Mean*	*SD*
Before-comparison				
Cognitive competence (out of 5)	*Verbal-first*	80	3.53	0.69
	Timeline-first	95	3.39	0.71
Interview performance (out of 4)	*Verbal-first*	79	2.80	0.53
	Timeline-first	92	2.77	0.54
Confidence (out of 5)	*Verbal-first*	80	2.99	0.96
	Timeline-first	95	3.07	1.03
Honesty (out of 5)	*Verbal-first*	80	3.73	0.82
	Timeline-first	95	3.74	0.86
Effort (out of 5)	*Verbal-first*	80	3.19	0.97
	Timeline-first	95	3.06	0.93
Suggestibility (out of 4)	*Verbal-first*	67	3.00	0.98
	Timeline-first	88	3.02	0.96
Cooperativeness (out of 5)	*Verbal-first*	80	4.23	0.86
	Timeline-first	95	4.02	0.92
Overall credibility (out of 5)	*Verbal-first*	80	3.35	0.84
	Timeline-first	95	3.27	0.85
After-comparison⁎				
Cognitive competence	*Verbal-first*	80	0.04	0.79
	Timeline-first	95	0.05	0.90
Interview performance	*Verbal-first*	80	0.01	0.85

		N	*Mean*	*SD*
	Timeline-first	95	-0.11	0.99
Confidence	*Verbal-first*	80	-0.01	0.91
	Timeline-first	94	-0.15	1.05
Honesty	*Verbal-first*	80	0.07	0.58
	Timeline-first	95	0.15	0.66
Effort	*Verbal-first*	80	-0.01	1.05
	Timeline-first	95	0.15	1.06
Suggestibility	*Verbal-first*	80	-0.09	0.86
	Timeline-first	95	-0.01	0.91
Cooperative	*Verbal-first*	80	-0.06	0.85
	Timeline-first	95	0.22	0.83
Overall credibility	*Verbal-first*	80	-0.01	0.82
	Timeline-first	95	-0.03	1.00
Timeline perceptions				
Helpfulness of the timeline	*Verbal-first*	80	2.69	1.09
	Timeline-first	95	3.12	1.16
Child's understanding of the timeline	*Verbal-first*	80	3.60	0.81
	Timeline-first	95	3.58	0.97
Overall impact of the timeline **	*Verbal-first*	80	0.38	0.90
	Timeline-first	95	0.59	0.78

Notes:

*After-comparison questions measured participants' perceptions of each dependent measure in comparison to the first interview video. Responses values range from -2 to 2, with negative

values representing a negative change, 0 representing “about the same”, and positive values representing a positive change.

** Overall impact ranged from -2 to 2, with negative values representing a negative change, 0 representing “no impact”, and positive values representing a positive change.

Table 20

Pearson Correlations Between Dependent Variables in the Before-Comparison Phase in Study 3

	1	2	3	4	5	6	7	8
1. Cognitive competence	-							
2. Interview performance	.61**	-						
3. Confidence	.52**	.56**	-					
4. Honesty	.53**	.58**	.59**	-				
5. Effort	.41**	.34**	.23**	.35**	-			
6. Suggestibility (Resistance)	.40**	.43**	.48**	.42**	.33**	-		
7. Cooperativeness	.35**	.50**	.37**	.41**	.36**	.39**	-	
8. Credibility	.66**	.68**	.70**	.78**	.38**	.50**	.40**	-

Note: *$p < .01$. **$p < .001$.

Table 21

Pearson Correlations Between Dependent Variables in the After-Comparison Phase in Study 3

	1	2	3	4	5	6	7	8
1. Cognitive competence	-							
2. Interview performance	.79**	-						
3. Confidence	.50**	.61**	-					
4. Honesty	.57**	.58**	.45**	-				
5. Effort	.35**	.28**	.10	.36**	-			
6. Suggestibility (Resistance)	.38**	.40**	.33**	.35**	.26**	-		
7. Cooperativeness	.25*	.38**	.29**	.39**	.37**	.27**	-	
8. Credibility	.73**	.83**	.65**	.69**	.36*	.42**	.40**	-

Note: *$p < .01$. **$p < .001$.

Table 22

Pearson Correlations Between Measures of Participants' Perceptions of the Timeline and the After-Comparison Credibility Rating in Study 3

	1	2	3	4
1. Helpfulness of the timeline	-			
2. Child's understanding of the timeline	.24**	-		
3. Overall impact of the timeline	.58**	.26*	-	
4. Overall credibility	.09	-.51	.06	-

Note: *$p < .01$. **$p < .001$.

Table 23

After-Comparison Overall Credibility Rating by Condition and Timeline Perceptions in Study 3.

		Verbal-first		*Timeline-first*	
	Timeline perceptions	Verbal video	Timeline video	Timeline video	Verbal video
Helpfulness of the timeline	Low		↓ *overall credibility*		↑ *overall credibility*
	High		↑ *overall credibility*		↓ *overall credibility*
Child's understanding of the timeline	Low		↓ *overall credibility*		↑ *overall credibility*
	High		↑ *overall credibility*		↓ *overall credibility*
Overall impact of the timeline	Negative		↓ *overall credibility*		↑ *overall credibility*
	Positive		↑ *overall credibility*		↓ *overall credibility*

Note: Low ratings in helpfulness of the timeline and the child's understanding of the timeline were represented by a rating of 1 standard deviation or more below the mean; high ratings in helpfulness of the timeline and the child's understanding of the timeline were represented by a rating of 1 standard deviation or more above the mean. Negative in overall impact of the timeline represented that the timeline was posing a negative impact on the child's recall; positive represented a perceived positive impact of the timeline.

References

Aldridge, J., Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Bowler, L. (2004). Using a human figure drawing to elicit information from alleged victims of child sexual abuse. *Journal of Consulting and Clinical Psychology, 72*(2), 304. doi:10.1037/0022-006X.72.2.304

Ames, L. B. (1946). The development of the sense of time in the young child. *The Pedagogical Seminary and Journal of Genetic Psychology, 68*(1), 97-125. doi:10.1080/08856559.1946.10533358

Anderson, J., Ellefson, J., Lashley, J., Miller, A. L., Olinger, S., & Russell, A. (2010). *The CornerHouse forensic interview protocol: RATAC®* (Vol. 12).

Bauer, P. J., Burch, M. M., Scholin, S. E., & Güler, O. E. (2007). Using cue words to investigate the distribution of autobiographical memories in childhood. *Psychological Science, 18*(10), 910-916. doi:10.1111/j.1467-9280.2007.01999.x

Bauer, P. J., & Mandler, J. M. (1992). Putting the horse before the cart: The use of temporal order in recall of events by one-year-old children. *Developmental Psychology, 28*(3), 441. doi:10.1037/0012-1649.28.3.441

Bottoms, B. L. (1993). Individual differences in perceptions of child sexual assault victims *Child victims, child witnesses: Understanding and improving testimony.* (pp. 229-261). New York, NY, US: Guilford Press.

Bottoms, B. L., & Goodman, G. S. (1994). Perceptions of children's credibility in sexual assault cases. *Journal of Applied Social Psychology, 24*(8), 702-732. doi:10.1111/j.1559-1816.1994.tb00608.x

Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review, 10*(1), 3-47. doi:10.1016/0273-2297(90)90003-M

Brainerd, C. J., & Reyna, V. F. (1998). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology, 71*(2), 81-129. doi:10.1006/jecp.1998.2464

Brainerd, C. J., & Reyna, V. F. (2004). Fuzzy-trace theory and memory development. *Developmental Review, 24*(4), 396-439. doi:10.1016/j.dr.2004.08.005

Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological bulletin, 134*(3), 343. doi:10.1037/0033-2909.134.3.343

Brainerd, C. J., Reyna, V. F., & Forrest, T. J. (2002). Are young children susceptible to the false-memory illusion? *Child Development, 73*(5), 1363-1377. doi:10.1111/1467-8624.00477

Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and human behavior, 26*(3), 353-364. doi:10.1023/A:1015380522722

Brewer, N., & Williams, K. D. (2005). *Psychology and law: An empirical perspective*: Guilford Publications.

Brown, D., Pipe, M.-E., Lewis, C., Lamb, M. E., & Orbach, Y. (2007). Supportive or suggestive: Do human figure drawings help 5-to 7-year old children to report touch? *Journal of Consulting and Clinical Psychology, 75*(1), 33-42. doi:10.1037/0022-006x.75.1.33

Brown, D., Pipe, M. E., Lewis, C., Lamb, M. E., & Orbach, Y. (2012). How Do Body Diagrams Affect the Accuracy and Consistency of Children's Reports of Bodily Touch Across Repeated Interviews? *Applied Cognitive Psychology, 26*(2), 174-181. doi:10.1002/acp.1828

Brown, G. D. A. (1997). Formal models of memory for serial order: A review. In M. A. Conway (Ed.), *Cognitive models of memory* (pp. 47-77). Cambridge, Massechusetts: The MIT Press.

Brubacher, S. P., Glisic, U., Roberts, K. P., & Powell, M. B. (2011). Children's ability to recall unique aspects of one occurrence of a repeated event. *Applied Cognitive Psychology, 25*(3), 351-358. doi:10.1002/acp.1696

Brubacher, S. P., Malloy, L. C., Lamb, M. E., & Roberts, K. P. (2013). How do interviewers and children discuss individual occurrences of alleged repeated abuse in forensic interviews? *Applied Cognitive Psychology, 27*(4), 443-450. doi:10.1002/acp.2920

Brubacher, S. P., Peterson, C., La Rooy, D., Dickinson, J. J., & Poole, D. A. (2019). How children talk about events: Implications for eliciting and analyzing eyewitness reports. *Developmental Review, 51*, 70-89. doi:10.1016/j.dr.2018.12.003

Brubacher, S. P., Roberts, K. P., & Powell, M. B. (2012). Retrieval of episodic versus generic information: Does the order of recall affect the amount and accuracy of details reported by children about repeated events? *Developmental Psychology, 48*(1), 111. doi:10.1037/a0025864

Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual review of psychology, 50*(1), 419-439. doi:10.1146/annurev.psych.50.1.419

Bruck, M., Ceci, S. J., & Francoeur, E. (2000). Children's use of anatomically detailed dolls to report genital touching in a medical examination: Developmental and gender comparisons. *Journal of Experimental Psychology-Applied, 6*(1), 74-83. doi:10.1037//0278-7393.6.1.74

Bruck, M., Francouer, E., Ceci, S. J., & Renick, A. (1995). Anatomically detailed dolls do not facilitate preschoolers reports of a pediatric examination Involving genital touching. *Journal of Experimental Psychology-Applied, 1*(2), 95-109. doi:10.1037//1076-898x.1.2.95

Bruck, M., Kelley, K., & Poole, D. A. (2016). Children's reports of body touching in medical examinations: The benefits and risks of using body diagrams. *Psychology, Public Policy, and Law, 22*(1), 1. doi:10.1037/law0000076

Bruer, K., & Pozzulo, J. D. (2014). Influence of eyewitness age and recall error on mock juror decision-making. *Legal and Criminological Psychology, 19*(2), 332-348. doi:10.1111/lcrp.12001

Bushnell, I. W. R. (2001). Mother's face recognition in newborn infants: learning and memory. *Infant and Child Development, 10*(1‐2), 67-74. doi:10.1002/icd.248

Castelli, P., Goodman, G. S., & Ghetti, S. (2005). Effects of Interview Style and Witness Age on Perceptions of Children's Credibility in Sexual Abuse Cases. *Journal of Applied Social Psychology, 35*(2), 297-317. doi:10.1111/j.1559-1816.2005.tb02122.x

Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: a historical review and synthesis. *Psychological bulletin, 113*(3), 403. doi:10.1037/0033-2909.113.3.403

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature, 395*(6699), 272-274. doi:10.1038/26216

Cleveland, K. C., Quas, J. A., & Lyon, T. D. (2016). Valence, implicated actor, and children's acquiescence to false suggestions. *Journal of Applied Developmental Psychology, 43*, 1-7. doi:10.1016/j.appdev.2015.12.003

Connolly, D. A., & Lindsay, D. S. (2001). The influence of suggestions on children's reports of a unique experience versus an instance of a repeated experience. *Applied Cognitive Psychology, 15*(2), 205-223. doi:10.1002/1099-0720(200103/04)15:2<205::AID-ACP698>3.0.CO;2-F

Connolly, D. A., Price, H., & Gordon, H. M. (2010). *Judicial decision making in timely and delayed prosecutions of child sexual abuse in Canada: A study of honesty and cognitive ability in assessments of credibility* (Vol. 16).

Connolly, D. A., & Price, H. L. (2006). Children's suggestibility for an instance of a repeated event versus a unique event: The effect of degree of association between variable details. *Journal of Experimental Child Psychology, 93*(3), 207-223. doi:10.1016/j.jecp.2005.06.004

Connolly, D. A., Price, H. L., Lavoie, J. A., & Gordon, H. M. (2008). Perceptions and predictors of children's credibility of a unique event and an instance of a repeated event. *Law and human behavior, 32*(1), 92. doi:10.1007/sl0979-006-9083-3

Cyr, M., & Lamb, M. E. (2009). Assessing the effectiveness of the NICHD investigative interview protocol when interviewing French-speaking alleged victims of child sexual abuse in Quebec. *Child Abuse & Neglect, 33*(5), 257-268. doi:10.1016/j.chiabu.2008.04.002

DeLoache, J. S. (1987). Rapid change in the symbolic functioning of very young children. *Science, 238*(4833), 1556-1557. doi:10.1126/science.2446392

DeLoache, J. S. (1991). Symbolic functioning in very young children: Understanding of pictures and models. *Child Development, 62*(4), 736-752. doi:10.2307/1131174

DeLoache, J. S. (2000). Dual representation and young children's use of scale models. *Child Development, 71*(2), 329-338. doi:10.1111/1467-8624.00148

DeLoache, J. S., & Marzolf, D. P. (1995). The use of dolls to interview young children: Issues of symbolic representation. *Journal of Experimental Child Psychology, 60*(1), 155-173. doi:10.1006/jecp.1995.1036

Earhart, B. (2017). *Young children's source monitoring: Exploring the contexts of task difficulty and repeated events.* (Unpublished doctoral dissertation), Wilfrid Laurier University, Waterloo, ON, Canada.

Earhart, B., & Roberts, K. P. (2014). The role of executive function in children's source monitoring with varying retrieval strategies. *Frontiers in Psychology, 5*(405). doi:10.3389/fpsyg.2014.00405

Eisen, M. L., Goodman, G. S., Qin, J., Davis, S., & Crayton, J. (2007). Maltreated children's memory: Accuracy, suggestibility, and psychopathology. *Developmental Psychology, 43*(6), 1275. doi:10.1037/0012-1649.43.6.1275

Evans, J. R., Michael, S. W., Meissner, C. A., & Brandon, S. E. (2013). Validating a new assessment method for deception detection: Introducing a Psychologically Based Credibility Assessment Tool. *Journal of Applied Research in Memory and Cognition, 2*(1), 33-41. doi:10.1016/j.jarmac.2013.02.002

Fallon, B., Van Wert, M., Trocmé, N., MacLaurin, B., Sinha, V., Lefebvre, R., . . . Rha, W. (2015). *Ontario incidence study of reported child abuse and neglect 2013*. Toronto, ON: Child Welfare Research Portal.

Farrar, M. J., & Boyer-Pennington, M. E. (1999). Remembering specific episodes of a scripted event. *Journal of Experimental Child Psychology, 73*, 266-288. doi:10.1006/jecp.1999.2507

Farrar, M. J., & Goodman, G. S. (1990). Developmental differences in the relation between scripts and episodic memory: Do they exist? *Knowing and remembering in young children.* (pp. 30-64). New York, NY, US: Cambridge University Press.

Farrar, M. J., & Goodman, G. S. (1992). Developmental changes in event memory. *Child Development, 63*(1), 173-187. doi:10.1111/j.1467-8624.1992.tb03605.x

Fawcett, H., & Winstanley, K. (2018). Children as alibi witnesses: the effect of age and confidence on mock-juror decision making. *Psychiatry, Psychology and Law, 25*(6), 957-971. doi:10.1080/13218719.2018.1482573

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., . . . Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), i-185. doi:10.2307/1166093

Field, T. M., Cohen, D., Garcia, R., & Greenberg, R. (1984). Mother-stranger face discrimination by the newborn. *Infant Behavior & Development, 7*(1), 19-25. doi:10.1016/S0163-6383(84)80019-3

Fisher, R., & Mitchell, G. (2009). The relation between consistency and accuracy of eyewitness testimony: Legal versus cognitive explanations. In R. Bull, T. Valentine, & T. Williamson (Eds.), *Handbook of Psychology of Investigative Interviewing* (pp. 121-136).

Fivush, R. (1984). Learning about School: The Development of Kindergartners' School Scripts. *Child Development, 55*(5), 1697-1709. doi:10.2307/1129917

Fivush, R., & Hamond, N. R. (1990). Autobiographical memory across the preschool years: Toward reconceptualizing childhood amnesia *Knowing and remembering in young children.* (pp. 223-248). New York, NY, US: Cambridge University Press.

Fivush, R., & Nelson, K. (2004). Culture and language in the emergence of autobiographical memory. *Psychological Science, 15*(9), 573-577. doi:10.1111/j.0956-7976.2004.00722.x

Friedman, W. J. (1986). The development of children's knowledge of temporal structure. *Child Development*, 1386-1400. doi:10.2307/1130418

Friedman, W. J. (1990). Children's representations of the pattern of daily activities. *Child Development, 61*(5), 1399-1412. doi:10.2307/1130751

Friedman, W. J. (1991). The development of children's memory for the time of past events. *Child Development, 62*(1), 139-155. doi:10.2307/1130710

Friedman, W. J. (1992). Children's time memory: The development of a differentiated past. *Cognitive Development, 7*(2), 171-187. doi:10.1016/0885-2014(92)90010-O

Friedman, W. J. (1993). Memory for the time of past events. *Psychological bulletin, 113*(1), 44-66. doi:10.1037/0033-2909.113.1.44

Friedman, W. J., Gardner, A. C., & Zubin, N. R. (1995). Children's comparisons of the recency of two events from the past year. *Child Development, 66*(4), 970-983. doi:10.1111/j.1467-8624.1995.tb00916.x

Friedman, W. J., & Kemp, S. (1998). The effects of elapsed time and retrieval on young children's judgments of the temporal distances of past events. *Cognitive Development, 13*(3), 335-367. doi:10.1016/S0885-2014(98)90015-6

Friedman, W. J., & Lyon, T. D. (2005). Development of temporal‐reconstructive abilities. *Child Development, 76*(6), 1202-1216. doi:10.1111/j.1467-8624.2005.00844.x-i1

Friedman, W. J., Reese, E., & Dai, X. (2011). Children's memory for the times of events from the past years. *Applied Cognitive Psychology, 25*(1), 156-165. doi:10.1002/acp.1656

Ghetti, S., Goodman, G. S., Eisen, M. L., Qin, J., & Davis, S. L. (2002). Consistency in children's reports of sexual and physical abuse. *Child Abuse & Neglect, 26*(9), 977-995. doi:10.1016/S0145-2134(02)00367-8

Goodman, G. S. (1980). Picture memory: How the action schema affects retention. *Cognitive Psychology, 12*(4), 473-495. doi:10.1016/0010-0285(80)90017-1

Goodman, G. S., Batterman-Faunce, J. M., Schaaf, J. M., & Kenney, R. (2002). Nearly 4 years after an event: Children's eyewitness memory and adults' perceptions of children's accuracy. *Child Abuse & Neglect, 26*(8), 849-884. doi:10.1016/S0145-2134(02)00354-X

Goodman, G. S., & Melinder, A. (2007). Child witness research and forensic interviews of young children: A review. *Legal and Criminological Psychology, 12*(1), 1-19. doi:10.1348/135532506X156620

Gosse, L. L., & Roberts, K. P. (2013). Children’s use of a ‘time line’to indicate when events occurred. *Journal of Police and Criminal Psychology, 29*(1), 36-43. doi:10.1007/s11896-013-9118-x

Granhag, P. A., & Strömwall, L. A. (2000). Effects of preconceptions on deception detection and new answers to why lie-catchers often fail. *Psychology, Crime & Law, 6*(3), 197-218. doi:10.1080/10683160008409804

Grant, J. B., & Suddendorf, T. (2009). Preschoolers begin to differentiate the times of events from throughout the lifespan. *European Journal of Developmental Psychology, 6*(6), 746-762. doi:10.1080/17405620802102947

Guadagno, B. L., Powell, M. B., & Wright, R. (2006). Police officers' and legal professionals' perceptions regarding how children are, and should be, questioned about repeated abuse. *Psychiatry, Psychology and Law, 13*(2), 251-260. doi:10.1375/pplt.13.2.251

Hamond, N. R., & Fivush, R. (1991). Memories of Mickey Mouse: Young children recount their trip to disneyworld. *Cognitive Development, 6*(4), 433-448. doi:10.1016/0885-2014(91)90048-I

Harner, L. (1975). Yesterday and tomorrow: Development of early understanding of the terms. *Developmental Psychology, 11*(6), 864-865. doi:10.1037/0012-1649.11.6.864

Hayne, H., & Rovee-Collier, C. (1995). The organization of reactivated memory in infancy. *Child Development, 66*(3), 893-906. doi:10.1111/j.1467-8624.1995.tb00912.x

Hayne, H., & Simcock, G. (2009). Memory development in toddlers. In M. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 43-68). New York: Psychology Press.

Higgs, W. L. (2013). *Using Images to Help Children Talk About Their Experiences.* (Unpublished Master's Thesis), Victoria University of Wellington, Victoria, BC, Canada.

Hintzman, D. L., Summers, J. J., & Block, R. A. (1975). What causes the spacing effect? Some effects of repetition, duration, and spacing on memory for pictures. *Memory & Cognition, 3*(3), 287-294. doi:10.3758/BF03212913

Hoodless, P. A. (2002). An investigation into children's developing awareness of time and chronology in story. *Journal of Curriculum Studies, 34*(2), 173-200. doi:10.1080/00220270110080962

Hope, L., Mullis, R., & Gabbert, F. (2013). Who? What? When? Using a timeline technique to facilitate recall of a complex event. *Journal of Applied Research in Memory and Cognition, 2*(1), 20-24. doi:10.1016/j.jarmac.2013.01.002

Howe, M. L., & Courage, M. (1993). On resolving the enigma of infantile amnesia. *Psychological bulletin, 113*(1), 305-326. doi:10.1037//0033-2909.113.2.305

Howe, M. L., & Courage, M. (1997). The emergence and early development of autobiographical memory. *Psychological Review, 104*(3), 499-523. doi:10.1037/0033-295X.104.3.499

Hudson, J. A., Fivush, R., & Kuebli, J. (1992). Scripts and episodes: The development of event memory. *Applied Cognitive Psychology, 6*(6), 483-505. doi:10.1002/acp.2350060604

Hudson, J. A., & Nelson, K. (1986). Repeated encounters of a similar kind: Effects of familiarity on children's autobiographic memory. *Cognitive Development, 1*(3), 253-271. doi:10.1016/S0885-2014(86)80004-1

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological bulletin, 114*(1), 3. doi:10.1037/0033-2909.114.1.3

Katz, C., & Hershkowitz, I. (2009). The effects of drawing on children’s accounts of sexual abuse. *Child Maltreatment, 15*(2), 171-179. doi:10.1177/1077559509351742

Kulkofsky, S., Wang, Q., & Ceci, S. J. (2008). Do better stories make better memories? Narrative quality and memory accuracy in preschool children. *Applied Cognitive Psychology, 22*(1), 21-38. doi:10.1002/acp.1326

La Rooy, D. J., Brubacher, S. P., Aromäki-Stratos, A., Cyr, M., Irit Hershkowitz, I., Korkman, J., . . . Roberts, K. P. (2015). The NICHD protocol: a review of an internationally-used evidence-based tool for training child forensic interviewers. *Journal of Criminological Research, Policy and Practice, 1*(2). doi:10.1108/JCRPP-01-2015-0001

Lamb, M. E., & Fauchier, A. (2001). The effects of question type on self-contradictions by children in the course of forensic interviews. *Applied Cognitive Psychology, 15*(5), 483-491. doi:10.1002/acp.726

Lamb, M. E., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2008). *Tell me what happened: Structured investigative interviews of child victims and witnesses*. Hoboken, NJ, US: John Wiley & Sons.

Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Boat, B., & Everson, M. D. (1996). Investigative interviews of alleged sexual abuse victims with and without anatomical dolls. *Child Abuse & Neglect, 20*(12), 1251-1259. doi:10.1016/S0145-2134(96)00121-4

Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., & Yudilevitch, L. (1996). Effects of investigative utterance types on Israeli children's responses. *International Journal of Behavioral Development, 19*(3), 627-637. doi:10.1177/016502549601900310

Lamb, M. E., Orbach, Y., Hershkowitz, I., Esplin, P. W., & Horowitz, D. (2007). A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: A review of research using the NICHD Investigative Interview Protocol. *Child Abuse & Neglect, 31*(11), 1201-1231. doi:10.1016/j.chiabu.2007.03.021

Leippe, M. R., Manion, A. P., & Romanczyk, A. (1993). Discernability or discrimination? Understanding jurors' reactions to accurate and inaccurate child and adult eyewitnesses. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving testimony.* (pp. 169-201). New York, NY, US: Guilford Press.

Lewandowsky, S., & Murdock Jr, B. B. (1989). Memory for serial order. *Psychological Review, 96*(1), 25-57. doi:10.1037/0033-295X.96.1.25

Linder, D. (2007). The McMartin preschool abuse trial. *Available at SSRN 1030559.*

Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology, 52*(3), 297-318. doi:10.1016/0022-0965(91)90065-Z

Lyon, T. D., & Saywitz, K. J. (2006). From post-mortem to preventive medicine: Next steps for research on child witnesses. *Journal of Social Issues, 62*(4), 833-861. doi:10.1111/j.1540-4560.2006.00489.x

Malloy, L. C. (2002). *Children's use of time line representations: Implications for investigative interviewing.* (Undergraduate Honours thesis), Central Michigan University, U.S.

McWilliams, K., Lyon, T. D., & Quas, J. A. (2016). Maltreated children's ability to make temporal judgments using a recurring landmark event. *Journal of Interpersonal Violence, 34*(4), 873-883. doi:10.1177/0886260516645812

Myers, J. E. B., Redlich, A. D., Goodman, G. S., Prizmich, L. P., & Imwinkelried, E. (1999). Jurors' perceptions of hearsay in child sexual abuse cases. *Psychology, Public Policy, and Law, 5*(2), 388-419. doi:10.1037/1076-8971.5.2.388

Neal, T. M. S., Christiansen, A., Bornstein, B. H., & Robicheaux, T. R. (2012). The effects of mock jurors' beliefs about eyewitness performance on trial judgments. *Psychology, Crime & Law, 18*(1), 49-64. doi:10.1080/1068316X.2011.587815

Nelson, K. (1988). The ontogeny of memory for real events. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory.* (pp. 244-276). New York, NY, US: Cambridge University Press.

Nelson, K. (1991). The matter of time: Interdependencies between language and thought in development. In J. P. Byrnes & S. A. Gelman (Eds.), *Perspectives on Language and*

Thought: Interrelations in Development (pp. 278-318). Cambridge: Cambridge University Press.

Nelson, K. (1993). The psychological and social origins of autobiographical memory. *Psychological Science, 4*(1), 7-14. doi:10.1111/j.1467-9280.1993.tb00548.x

Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: a social cultural developmental theory. *Psychological Review, 111*(2), 486-511. doi:10.1037/0033-295X.111.2.486

Nelson, K., Fivush, R., Hudson, J., & Lucariello, J. (1983). Scripts and the development of memory. In M. T. H. Chi (Ed.), *Trends in memory development research* (Vol. 9, pp. 52-70). Pittsburgh, PA, US: Karger Publishers.

Newcombe, P. A., & Bransgrove, J. (2007). Perceptions of witness credibility: Variations across age. *Journal of Applied Developmental Psychology, 28*(4), 318-331. doi:10.1016/j.appdev.2007.04.003

Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse & Neglect, 24*(6), 733-752. doi:10.1016/S0145-2134(00)00137-X

Otgaar, H., van Ansem, R., Pauw, C., & Horselenberg, R. (2016). Improving children's interviewing methods? The effects of drawing and practice on children's memories for an event. *Journal of Police and Criminal Psychology, 31*(4), 279-287. doi:10.1007/s11896-016-9190-0

Pawlak, A., Oehlrich, J. S., & Weist, R. M. (2006). Reference time in child English and Polish. *First Language, 26*(3), 281-297. doi:10.1177/0142723706059447

Pearse, S. L., Powell, M. B., & Thomson, D. M. (2003). The effect of contextual cues on children's ability to remember an occurrence of a repeated event. *Legal and Criminological Psychology, 8*(1), 39-50. doi:10.1348/135532503762871228

Peterson, C. C. (1996). Perceived truthfulness of children's and adults' testimony: The oath versus the competency test. In G. Davies, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology, law and criminal justice: International Developments in Research and Practice* (pp. 13-20). New York: de Gruyter.

Pezdek, K., & Roe, C. (1995). The effect of memory trace strength on suggestibility. *Journal of Experimental Child Psychology, 60*(1), 116-128. doi:10.1006/jecp.1995.1034

Pipe, M.-E., Thierry, K. L., & Lamb, M. E. (2007). The development of event memory: Implications for child witness testimony *The handbook of eyewitness psychology, Vol I: Memory for events.* (pp. 453-478). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Poole, D. A., & Bruck, M. (2012). Divining testimony? The impact of interviewing props on children's reports of touching. *Developmental Review, 32*(3), 165-180. doi:10.1016/J.Dr.2012.06.007

Poole, D. A., Bruck, M., & Pipe, M.-E. (2011). Forensic interviewing aids: Do props help children answer questions about touching? *Current Directions in Psychological Science, 20*(1), 11-15. doi:10.1177/0963721410388804

Poole, D. A., & Dickinson, J. J. (2011). Evidence supporting restrictions on uses of body diagrams in forensic interviews. *Child Abuse & Neglect, 35*(9), 659-669. doi:10.1016/j.chiabu.2011.05.004

Powell, M. B., Fisher, R. P., & Wright, R. (2005). Investigative interviewing. In N. Brewer & K. D. Williams (Eds.), *Psychology and law: An empirical perspective.* (pp. 11-42). New York, NY, US: The Guilford Press.

Powell, M. B., & Roberts, K. P. (2002). The effect of repeated experience on children's suggestibility across two question types. *Applied Cognitive Psychology, 16*(4), 367-386. doi:10.1002/acp.801

Powell, M. B., Roberts, K. P., Ceci, S. J., & Hembrooke, H. (1999). The effects of repeated experience on children's suggestibility. *Developmental Psychology, 35*(6), 1462. doi:10.1037/0012-1649.35.6.1462

Powell, M. B., Roberts, K. P., & Thomson, D. M. (2000). The effect of a suggestive interview on children's memory of a repeated event: Does it matter whether suggestions are linked to a particular incident? *Psychiatry, Psychology and Law, 7*(2), 182-191. doi:10.1080/13218710009524984

Powell, M. B., & Thomson, D. M. (1996). Children's memory of an occurrence of a repeated event: Effects of age, repetition, and retention interval across three question types. *Child Development, 67*(5), 1988-2004. doi:10.2307/1131605

Pozzulo, J. D., Lemieux, J. M. T., Wells, E., & McCuaig, H. J. (2006). The influence of eyewitness identification decisions and age of witness on jurors' verdicts and perceptions of reliability. *Psychology, Crime & Law, 12*(6), 641-652. doi:10.1080/10683160500415539

Quas, J. A., Thompson, W. C., Alison, K., & Stewart, C. (2005). Do jurors "know" what isn't so about child witnesses? *Law and human behavior, 29*(4), 425-456. doi:10.1007/s10979-005-5523-8

Roberts, K. P., & Blades, M. (1998). The effects of interacting in repeated events on children's eyewitness memory and source monitoring. *Applied Cognitive Psychology, 12*(5), 489-503. doi:10.1002/(SICI)1099-0720(199810)12:5<489::AID-ACP535>3.0.CO;2-%23

Ross, D. F., Dunning, D., Toglia, M. P., & Ceci, S. J. (1990). The child in the eyes of the jury: Assessing mock jurors' perceptions of the child witness. *Law and human behavior, 14*(1), 5-23. doi:10.1007/BF01055786

Ross, D. F., Miller, B. S., & Moran, P. B. (1987). The child in the eyes of the jury: Assessing mock jurors' perceptions of the child witness. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's Eyewitness Memory* (pp. 142-154). New York, NY: Springer US.

Rovee, C. K., & Rovee, D. T. (1969). Conjugate reinforcement of infant exploratory behavior. *Journal of Experimental Child Psychology, 8*(1), 33-39. doi:10.1016/0022-0965(69)90025-3

Salmon, K., Pipe, M. E., Malloy, A., & Mackay, K. (2012). Do non‐verbal aids increase the effectiveness of 'Best Practice' verbal interview techniques? An experimental study. *Applied Cognitive Psychology, 26*(3), 370-380. doi:10.1002/acp.1835

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Oxford, England: Lawrence Erlbaum.

Schmidt, C. W., & Brigham, J. C. (1996). Jurors' perceptions of child victim-witnesses in a simulated sexual abuse trial. *Law and human behavior, 20*(6), 581-606. doi:10.1007/BF01499233

Schneider, L., Price, H. L., Roberts, K. P., & Hedrick, A. M. (2011). Children's episodic and generic reports of alleged abuse. *Applied Cognitive Psychology, 25*(6), 862-870. doi:10.1002/acp.1759

Shaw, J. S., & McClure, K. A. (1996). Repeated postevent questioning can lead to elevated levels of eyewitness confidence. *Law and human behavior, 20*(6), 629-653. doi:10.1007/BF01499235

Simcock, G., & Hayne, H. (2002). Breaking the barrier? children fail to translate their preverbal memories into language. *Psychological Science, 13*(3), 225-231. doi:10.1111/1467-9280.00442

Slackman, E. A., Hudson, J. A., & Fivush, R. (1986). Actions, actors, links, and goals: The structure of children's event representations. In K. Nelson (Ed.), *Event knowledge: Structure and function in development* (pp. 47-69). Hillsdale, NJ, US: Lawrence Erlbaum Associates Publishers.

Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Esplin, P. W., Redlich, A., & Sunshine, N. (1996). The relation between investigative utterance types and the informativeness of child witnesses. *Journal of Applied Developmental Psychology, 17*(3), 439-451. doi:10.1016/S0193-3973(96)90036-2

Sternberg, K. J., Lamb, M. E., Orbach, Y., Esplin, P. W., & Mitchell, S. (2001). Use of a structured investigative protocol enhances young children's responses to free-recall prompts in the course of forensic interviews. *Journal of applied psychology, 86*(5), 997-1005. doi:10.1037/0021-9010.86.5.997

Teoh, Y.-S., Yang, P. J., Lamb, M. E., & Larsson, A. S. (2010). Do human figure diagrams help alleged victims of sexual abuse provide elaborate and clear accounts of physical contact with alleged perpetrators? *Applied Cognitive Psychology, 24*(2), 287-300. doi:10.1002/acp.1564

Thierry, K. L., Lamb, M. E., Orbach, Y., & Pipe, M.-E. (2005). Developmental differences in the function and use of anatomical dolls during interviews with alleged sexual abuse victims. *Journal of Consulting and Clinical Psychology, 73*(6), 1125. doi:10.1037/0022-006X.73.6.1125

Tillman, K. A., Marghetis, T., Barner, D., & Srinivasan, M. (2017). Today is tomorrow's yesterday: Children's acquisition of deictic time words. *Cognitive Psychology, 92*, 87-100. doi:10.1016/j.cogpsych.2016.10.003

Tulving, E. (2002). Episodic memory: from mind to brain. *Annual review of psychology, 53*(1), 1-25. doi:10.1146/annurev.psych.53.100901.135114

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352-373. doi:10.1037/h0020071

U.S. v. Tsinhnahijinnie. 112 F. 3d 988 (9th Cir. 1997).

Van Der Vaart, W., & Glasner, T. (2007). Applying a timeline as a recall aid in a telephone survey: A record check study. *Applied Cognitive Psychology, 21*(2), 227-238. doi:10.1002/acp.1338

Vredeveldt, A., & Sauer, J. (2015). Effects of eye-closure on confidence-accuracy relations in eyewitness testimony. *Journal of Applied Research in Memory and Cognition, 4*(1), 51-58. doi:10.1016/j.jarmac.2014.12.006

Weinsheimer, C. C., Coburn, P. I., Chong, K., MacLean, C. L., & Connolly, D. A. (2017). Perceptions of credibility for a memory report of a single versus repeated event. *Applied Cognitive Psychology, 31*(4), 414-423. doi:10.1002/acp.3340

Whitcomb, D. (1991). Improving the investigation and prosecution of child sexual-abuse cases: Research findings, questions, and implications for public policy. In D. D. Knudsen & J. L.

Miller (Eds.), *Abused and battered: Social and legal responses to family violence* (pp. 181-190). New York, NY, US: Aldine de Gruyter.

Willcock, E., Morgan, K., & Hayne, H. (2006). Body maps do not facilitate children's reports of touch. *Applied Cognitive Psychology, 20*(5), 607-615. doi:10.1002/acp.1212

Woiwod, D. M., Fitzgerald, R. J., Sheahan, C. L., Price, H. L., & Connolly, D. A. (2019). A meta-analysis of differences in children's reports of single and repeated events. *Law and human behavior, 43*(1), 99. doi:10.1037/lhb0000312

Wolfman, M., Brown, D., & Jose, P. (2018). The use of visual aids in forensic interviews with children. *Journal of Applied Research in Memory and Cognition, 7*(4), 587-596. doi:10.1016/j.jarmac.2018.06.004

World Health Organization. (2003). Guidelines for medico-legal care of victims of sexual violence.

Zhang, M., & Hudson, J. A. (2018). Children's understanding of yesterday and tomorrow. *Journal of Experimental Child Psychology, 170*, 107-133. doi:10.1016/j.jecp.2018.01.010

Figures

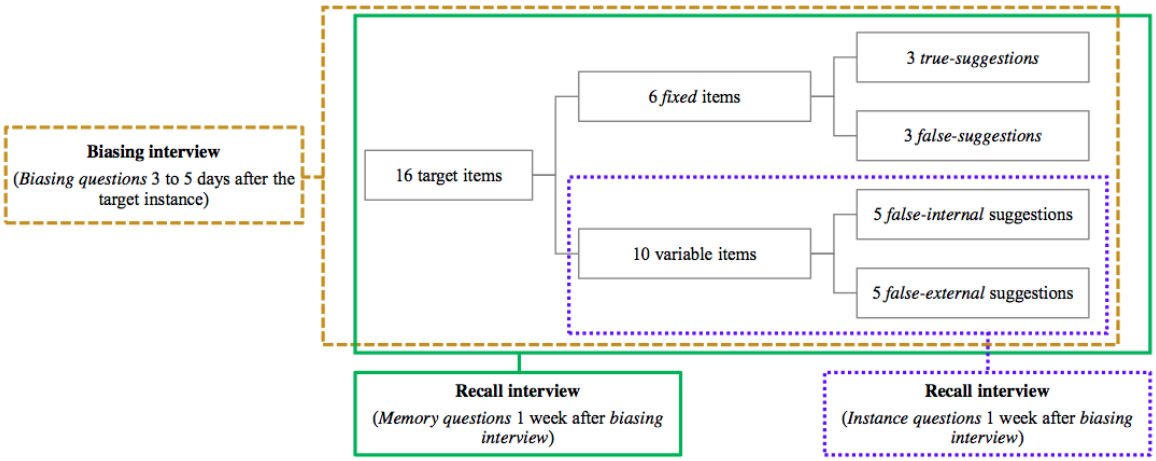


Figure 1: Interview structure and suggestion types for Study 1.

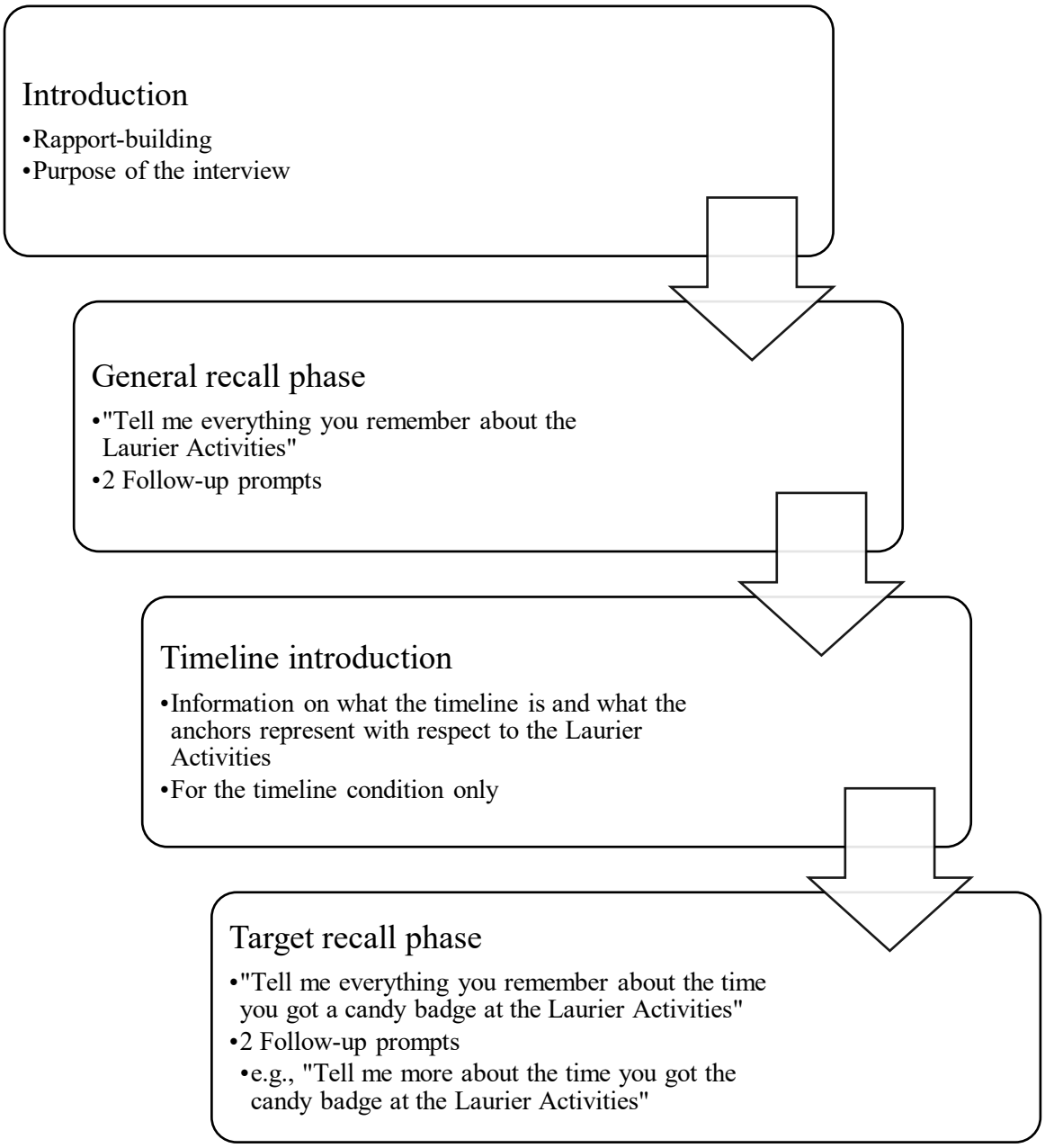


Figure 2: Interview procedure for Study 2

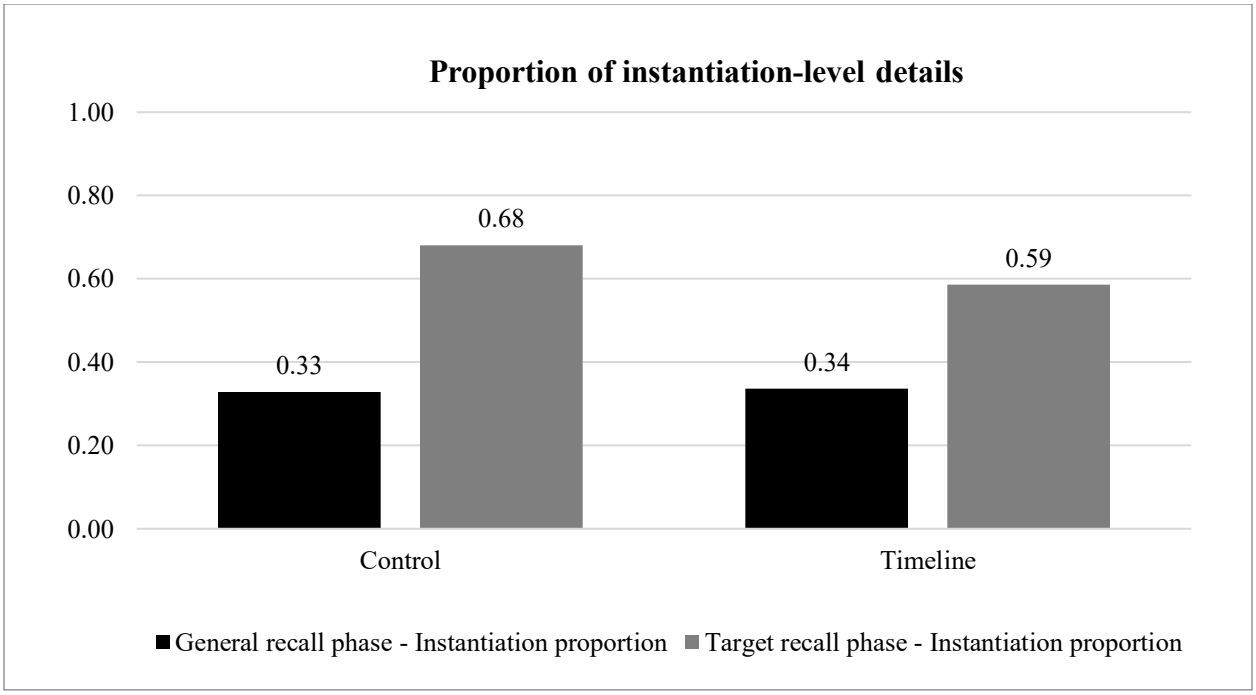


Figure 3: The proportion of instantiation-level details in each recall phase by interview condition for Study 2.

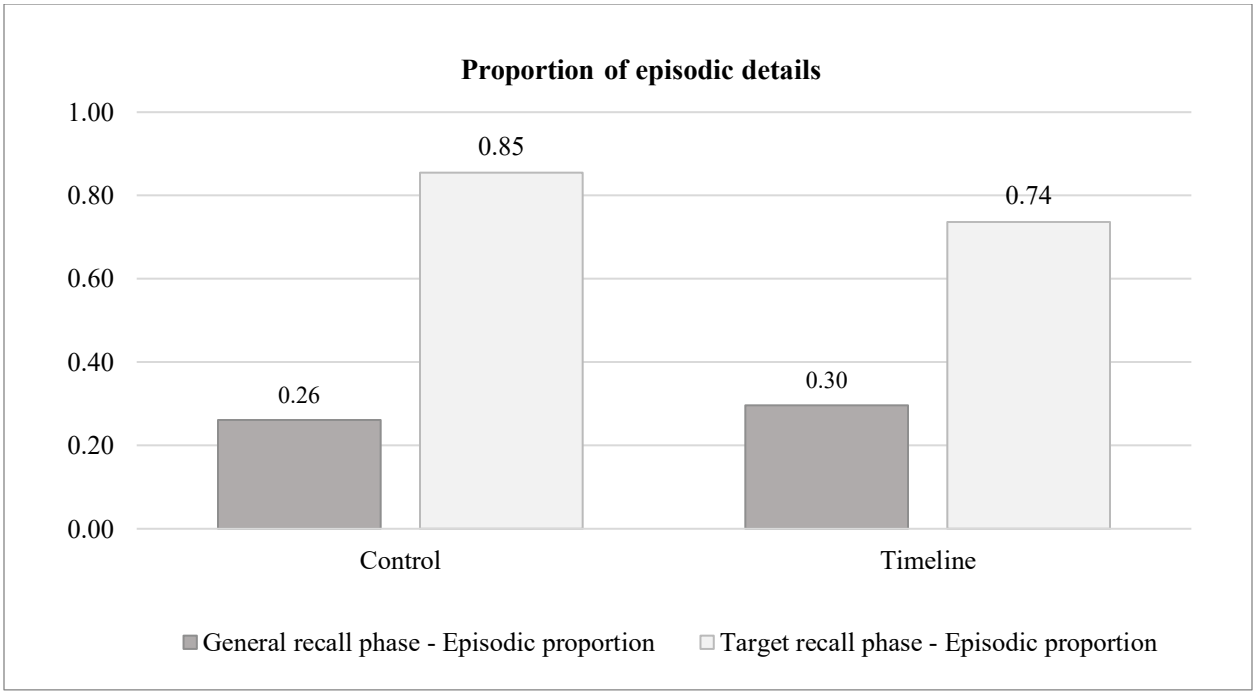


Figure 4: The proportion of episodic details in each recall phase by interview condition for Study 2.

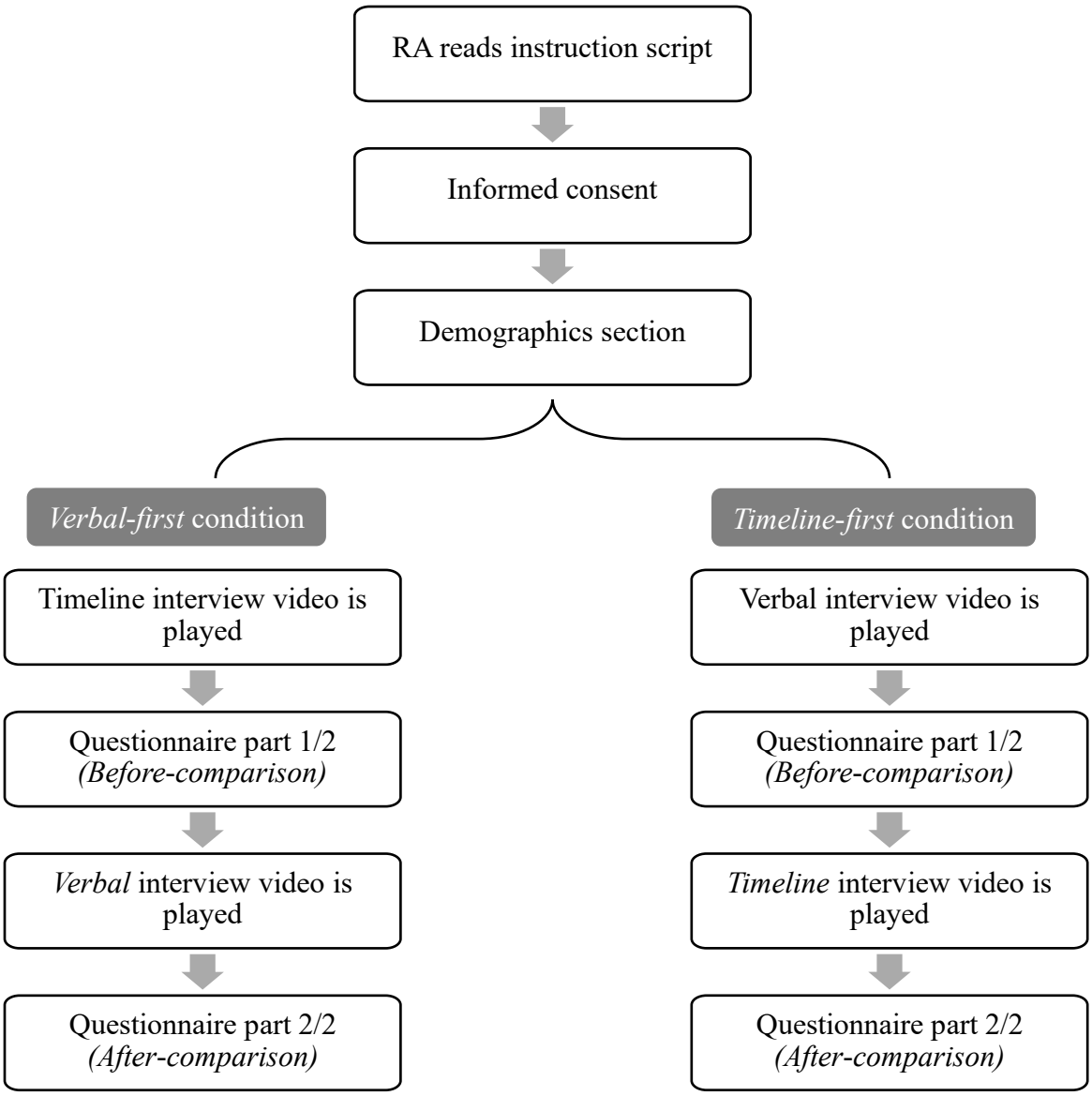


Figure 5: Study 3 procedure.

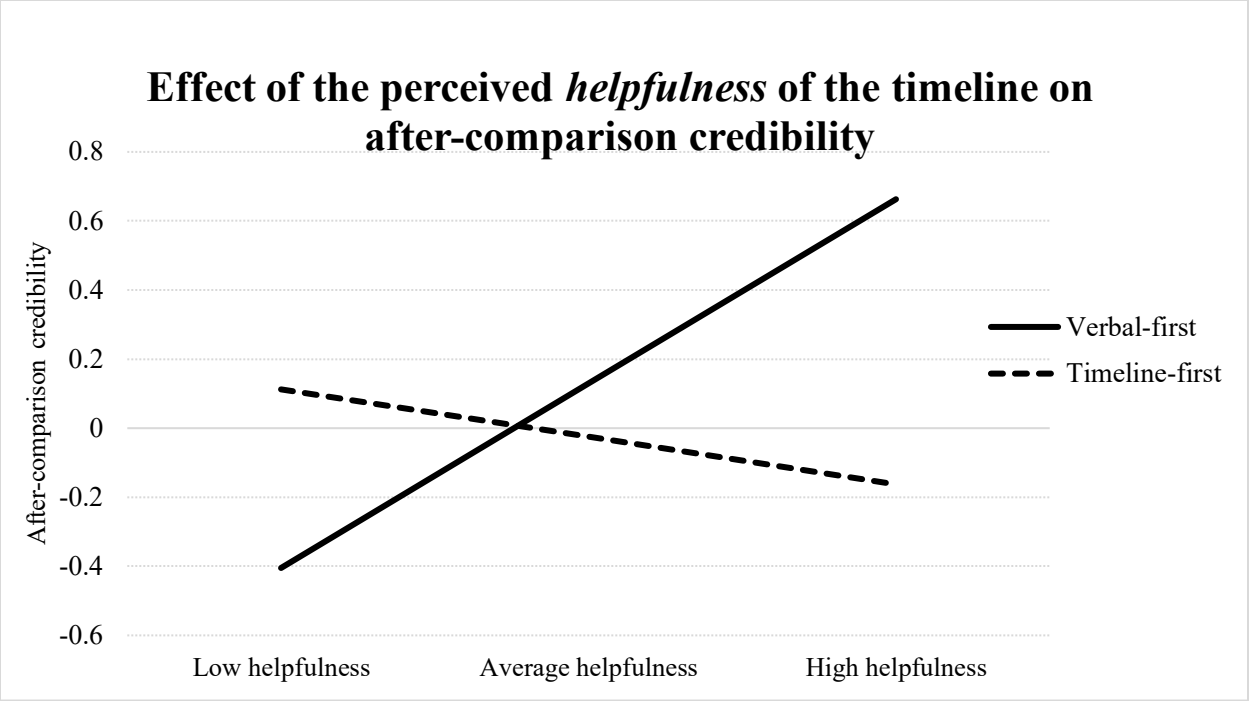


Figure 6: Moderation effect of perceived helpfulness of the timeline on after-comparison credibility for Study 3

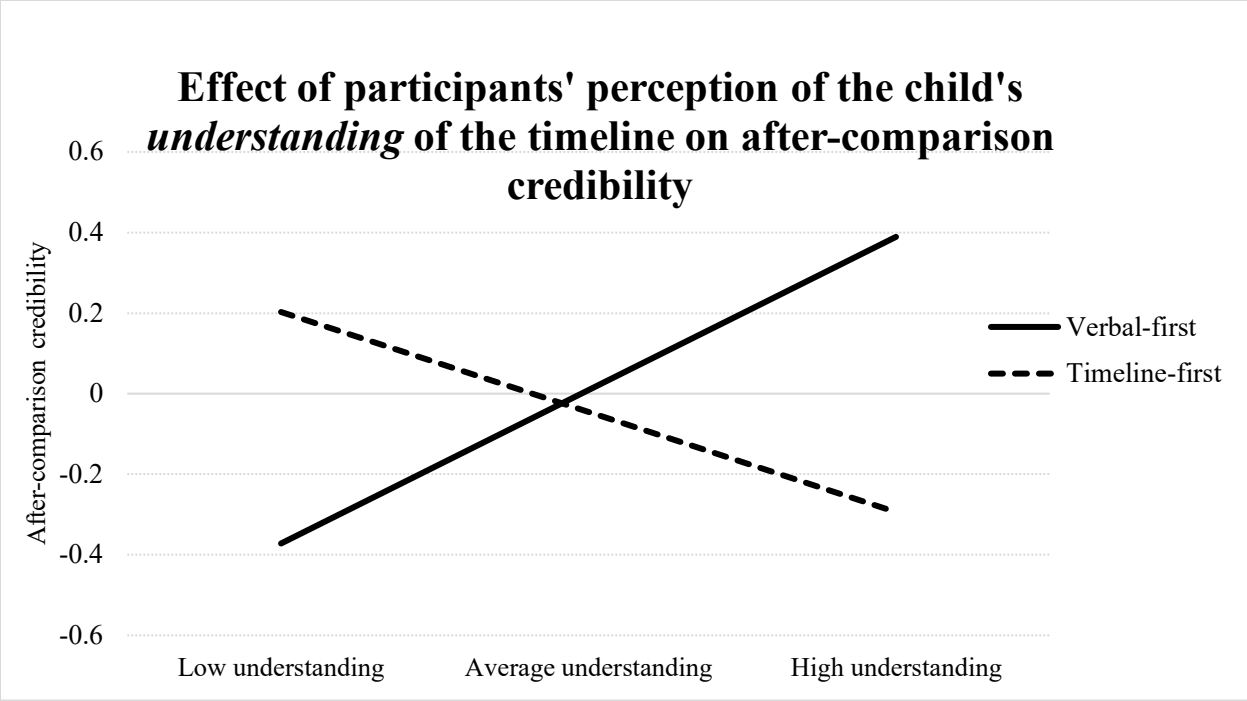


Figure 7: Moderation effect of perceived timeline understanding on after-comparison credibility for Study 3

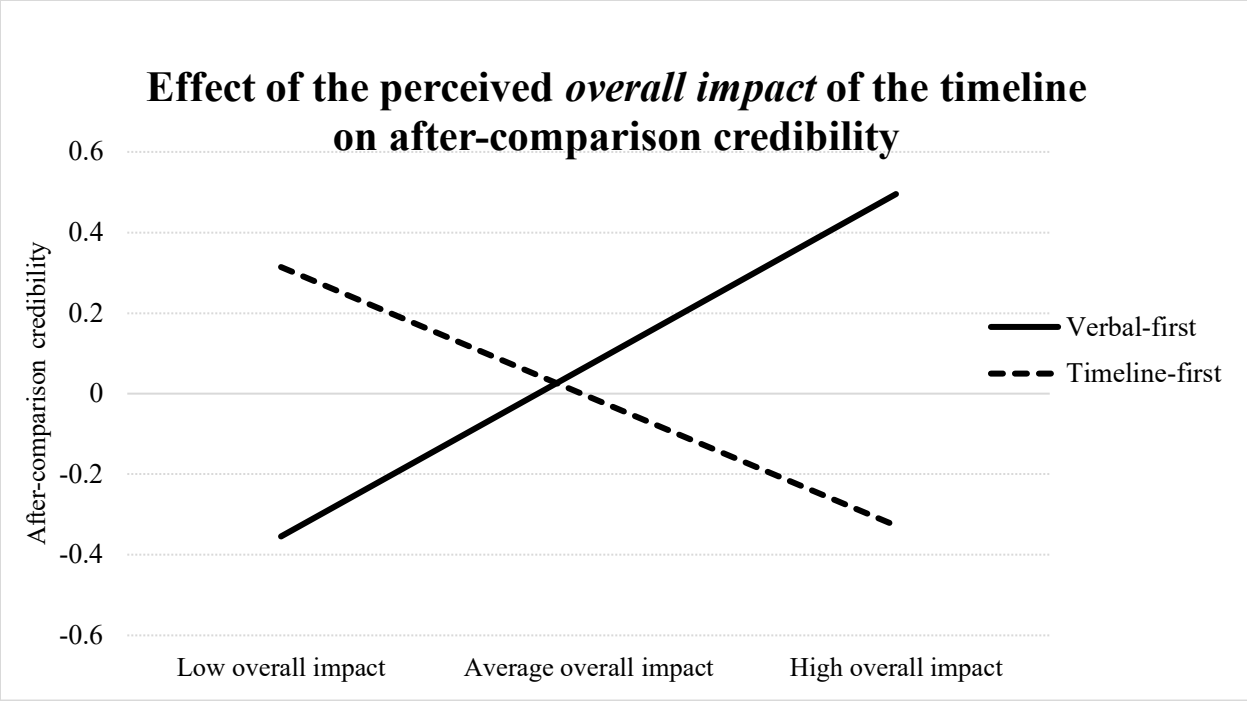


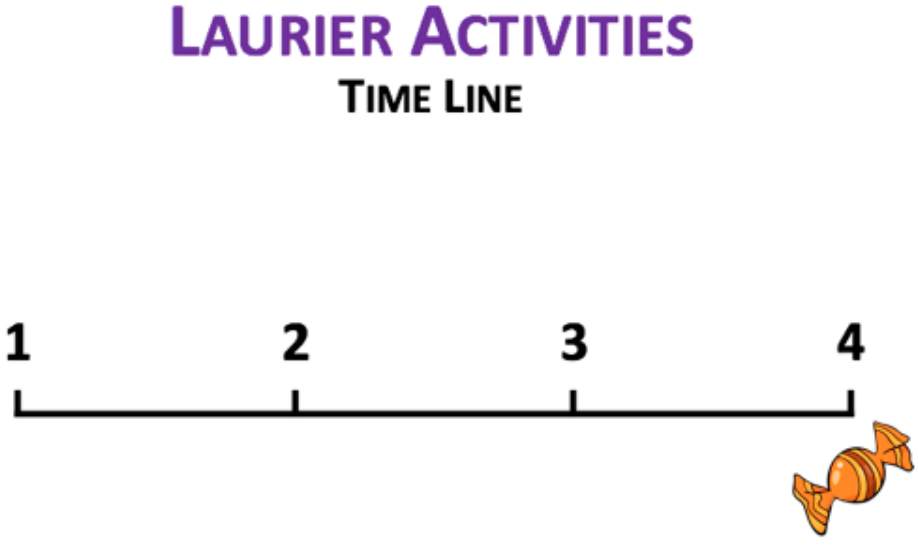
Figure 8: Moderation effect of perceived overall impact of the timeline on after-comparison credibility for Study 3

Appendices

Appendix A

Timelines used in Study 1 and 2

(Study 1)



(Study 2)

Appendix B

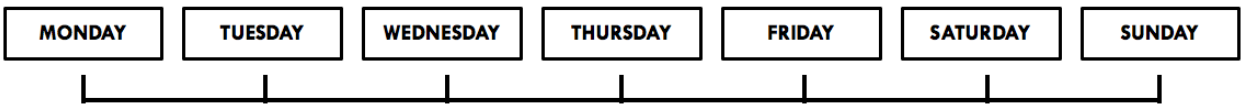
Interview questions used in child interview videos for Study 3

1. On which day or days of the week did you *go play outside* at BrainWorx?
2. On which day or days of the week did you learn about *your stomach and intestine*?
3. On which day or days of the week did you *go to a water park* at BrainWorx?
4. On which day or days of the week did you *learn how to fold paper swans* at BrainWorx?
5. On which day or days of the week did you *play scrabble board game* at BrainWorx?
6. I know you got to watch movies at BrainWorx. On which day or days of the week did you watch *Charlie and the Chocolate Factory* at BrainWorx?
7. On which day or days of the week did you *go rock climbing* at BrainWorx?
8. On which day or days of the week did you *go to a room with lots of red or green balloons* at BrainWorx?
9. There were different themes at BrainWorx, on which day or days of the week was the theme about *senses*?
10. There were different themes at BrainWorx, on which day or days of the week was the theme about *reflexes and reactions*?
11. There were different themes at BrainWorx, on which day or days of the week was the theme about *brain protection*?
12. I know you got to watch movies at BrainWorx. On which day or days of the week did you watch *Lorax* at BrainWorx?
13. I know you got to watch movies at BrainWorx. On which day or days of the week did you watch *Flubber* at BrainWorx?

14. I know you got to watch movies at BrainWorx. On which day or days of the week did you watch *Megamind* at BrainWorx?
15. There were science experiments at BrainWorx. On which day or days of the week did you do a *slime experiment* at BrainWorx?
16. There were science experiments at BrainWorx. On which day or days of the week did you do a *kneejerk experiment* at BrainWorx?

Appendix C

Timeline used in Study 3



Appendix D

Full questionnaires used in Study 3 [1]

[1] Note: Participants in both conditions (*verbal first* and *timeline first*) responded to the same set of questions. However, the difference between the two conditions was that the Timeline section of the questionnaire always followed after participants have watched the timeline interview. In other words, in the *verbal first* condition, the Timeline section followed after Video 2; in the *timeline first* condition, the Timeline section followed after Video 1.

Demographics section

B01 1. Gender (Leave blank if you prefer not to answer): _____

B02 Age (in years) _____

B03 Program of Study

▼ Ancient Studies (1) ... OTHER (61)

B04 Year of program you are currently in

▼ 1st year (1) ... 4th year or above (4)

B05 Ethnicity

- White (1)
- Black or African Canadian (2)
- Indigenous (3)
- East Asian (4)
- South Asian (5)
- Hispanic or Latino (6)
- Other (8)

B06 How would you describe your experience with children aged between 6 and 9 years old?

- No experience at all (1)
- Observational experience only (i.e., experience where no interaction took) (2)
- Little interactive experience (3)
- Moderate interactive experience (4)
- Lots of interactive experience (5)

Video 1

Please answer the following questions regarding the interview video you just watched.

C01 What is the age of the child in the interview video?

- Under 7 years old (1)
- Between 7 and 8 years old (2)
- Between 8 and 9 years old (3)
- 9 years or older (4)
- I don't know (5)

C02 How intelligent do you think the child was?

- Not intelligent at all (1)
- Slightly intelligent (2)
- Moderately intelligent (3)
- Intelligent (4)
- Very intelligent (5)

C03 How difficult do you believe the questions were to the child?

- Extremely easy (1)
- Somewhat easy (2)
- Neither easy nor difficult (3)
- Somewhat difficult (4)
- Extremely difficult (5)

C04 How do you evaluate the child's understanding of days in the week as demonstrated in this interview?

- No understanding at all (1)
- Little understanding (2)
- Some understanding (3)
- Good understanding (4)
- Very good understanding (5)

D01 How accurate do you think the child was throughout the interview?

- Not accurate at all (1)
- Mostly inaccurate (2)
- I don't know (3)
- Mostly accurate (4)
- Very accurate (5)

D02 How coherent do you think the child was throughout the interview?

- Not coherent at all (1)
- Mostly in coherent (2)
- I don't know (3)
- Mostly coherent (4)
- Very coherent (5)

D03 How do you evaluate the length of time the child took in answering the interviewer’s questions?

- Very slow at answering (1)
- Somewhat slow at answering (2)
- Just right (3)
- Answered somewhat quickly (4)
- Answered very quickly (5)

D04 How does the length of time the child took in answering the questions affect your perception of his/her performance?

- Very negatively (1)
- Somewhat negatively (2)
- No effect (3)
- Somewhat positively (4)
- Very positively (5)

D05 How inaccurate do you believe the child was throughout the interview? ?

- Very inaccurate (1)
- Somewhat inaccurate (2)
- Somewhat accurate (3)
- Very accurate (4)
- I don't know (5)

D06 How do you describe the child's response to misleading questions?

- Easily misled (1)
- Somewhat misled (2)
- Somewhat resistant to misleading questions (3)
- Mostly resistant to misleading questions (4)
- I don' t know (5)

D07 How much effort did the child put into answering the questions?

- No effort at all (1)
- Little effort (2)
- Some effort (3)
- A moderate amount of effort (4)
- A great deal of effort (5)

D08 How confident do you think the child was during the interview?

- Not at all confident (1)
- Slightly confident (2)
- Moderately confident (3)
- Confident most of the time (4)
- Extremely confident (5)

D09 How cooperative do you think the child was during the interview?

- Not cooperative at all (1)
- Slightly cooperative (2)
- Moderately cooperative (3)
- Cooperative (4)
- Very cooperative (5)

E01 How honest do you think the child was?

- Not at all honest (1)
- Slightly honest (2)
- Moderately honest (3)
- Honest most of the time (4)
- Extremely honest (5)

E02 How believable do you think the child was?

- Not at all believable (1)
- Slightly believable (2)
- Moderately believable (3)
- Believable most of the time (4)
- Extremely believable (5)

E03 How truthful do you think the child was?

- Not at all truthful (1)
- Slightly truthful (2)
- Moderately truthful (3)
- Truthful most of the time (4)
- Extremely truthful (5)

E04 Overall, using a scale of 1(least credible) to 5 (very credible) how credible do you think the child's account of BrainWorx was?

- 1-Least credible (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5-Very credible (5)

Timeline section

This section of the questionnaire followed the *video 1 section* of the questionnaire in the *timeline first* condition; and the *video 2 section* of the questionnaire in the *verbal first* condition.

G01 How well do you think the child understood the timeline?

- No understanding at all (1)
- Little understanding (2)
- Some understanding (3)
- Good understanding (4)
- Very good understanding (5)

G02 How helpful do you think the timeline was to the child?

- Not at all helpful (1)
- Slightly helpful (2)
- Moderately helpful (3)
- Helpful (4)
- Extremely helpful (5)

G03 How do you think the timeline affected the child's accuracy?

- Very negatively (1)
- Slightly negatively (2)
- No effect (3)
- Slightly positively (4)
- Very positively (5)

G04 How did the timeline affect the child's ability to answer the questions?

- Very negatively (1)
- Slightly negatively (2)
- No effect (3)
- Slightly positively (4)
- Very positively (5)

G05 Based on your observation, how often did the child answer verbally AND demonstrated using the timeline (i.e., answered verbally as well as pointed to the day(s) on the timeline?

- Almost always (1)
- Often (2)
- Sometimes (3)
- Seldom (4)
- Never (5)
- I am not sure (6)

Video 2

Please answer the following questions regarding both interview videos you watched today.

F01 How intelligent do you think the child was, compared to what you saw in the first interview?

- A lot less intelligent (1)
- Slightly less intelligent (2)
- About the same (3)
- Slightly more intelligent (4)
- A lot more intelligent (5)

F02 How difficult do you believe the questions were to the child, compared to what you saw in the first interview?

- A lot easier (1)
- Slightly easier (2)
- About the same (3)
- Slightly more difficult (4)
- A lot more difficult (5)

F03 How do you evaluate the child's understanding of days in the week demonstrated in this interview, compared to what you saw in the first video?

- A lot less understanding (1)
- Slightly less understanding (2)
- About the same (3)
- Slightly more understanding (4)
- A lot more understanding (5)

F04 How accurate do you think the child was throughout this interview, compared to what you saw in the first interview?

- A lot less accurate (1)
- Slightly less accurate (2)
- About the same (3)
- Slightly more accurate (4)
- A lot more accurate (5)

F05 How coherent do you think the child was throughout this interview, compared to what you saw in the first interview?

- A lot less coherent (1)
- Slightly less coherent (2)
- About the same (3)
- Slightly more coherent (4)
- A lot more coherent (5)

F06 How do you evaluate the length of time the child took in answering the interviewer's questions, compared to what you saw in the first interview?

- A lot slower at answering (1)
- Slightly slower at answering (2)
- About the same (3)
- Slightly quicker at answering (4)
- A lot quicker at answering (5)

F07 Compared to the first interview, the length of time the child took in answering the questions affected my perceptions of his/her performance _____.

- A lot less significantly in this interview (1)
- Slightly less significantly in this interview (2)
- About the same (3)
- Slightly more significantly in this interview (4)
- A lot more significantly in this interview (5)

F08 How inaccurate do you think the child was in this interview, compared to what you saw in the first interview?

- A lot less inaccurate (1)
- Slightly less inaccurate (2)
- About the same (3)
- Slightly more inaccurate (4)
- A lot more inaccurate (5)

F09 How do you describe the child's response to misleading questions in this interview, compared to what you saw in the first interview?

- A lot more easily misled (1)
- Slightly more misled (2)
- About the same (3)
- Slightly more resistant to misleading questions (4)
- A lot more resistant to misleading questions (5)

F10 How much effort do you think the child put into answering the questions in this interview, compared to what you saw in the first interview?

- A lot less effort (1)
- Slightly less effort (2)
- About the same (3)
- Slightly more effort (4)
- A lot more effort (5)

F11 How confident do you think the child was in this interview, compared to what you saw in the first interview?

- A lot less confident (1)
- Slightly less confident (2)
- About the same (3)
- Slightly more confident (4)
- A lot more confident (5)

F12 How cooperative do you think the child was in this interview, compared to what you saw in the first interview?

- A lot less cooperative (1)
- Slightly less cooperative (2)
- About the same (3)
- Slightly more cooperative (4)
- A lot more cooperative (5)

H01 How honest do you think the child was, compared to what you saw in the first interview?

- A lot less honest (1)
- Slightly less honest (2)
- About the same (3)
- Slightly more honest (4)
- A lot more honest (5)

H02 How believable do you think the child was, compared to what you saw in the first interview?

- A lot less believable (1)
- Slightly less believable (2)
- About the same (3)
- Slightly more believable (4)
- A lot more believable (5)

H03 How truthful do you think the child was, compared to what you saw in the first interview?

- A lot less truthful (1)
- Slightly less truthful (2)
- About the same (3)
- Slightly more truthful (4)
- A lot more truthful (5)

H04 How likeable do you think the child was, compared to what you saw in the first interview?

- A lot less likeable (1)
- Slightly less likeable (2)
- About the same (3)
- Slightly more likeable (4)
- A lot more likeable (5)

H05 Overall, how credible was the child's account of BrainWorx, compared to what you saw in the first interview?

- A lot less credible (1)
- Slightly less credible (2)
- About the same (3)
- Slightly more credible (4)
- A lot more credible (5)

Appendix E

Instruction script for Study 3

You are about to watch videos of a child being interviewed about the Laurier BrainWorx summer camp. The child participated in the camp from Monday to Friday during a week over the summer of 2017. A series of activities were conducted, some happened once, some happened every day, some happened more than once. The child had participated in the first three days of the camp on Monday, Tuesday, and Wednesday and was interviewed on Thursday. Questions were asked about activities that happened or did not happen. The interviewer in the video followed strictly to a set of interview questions and did not know the correct answers. The interviewer may be seen repeating the child's response, confirming an answer from the child, and/or saying "mm-hmm"; these were simply natural behaviors or behaviors encouraging the child to participate in the questioning process. Please do not misconstrue these as signs of correcting the child's response or account them into your decision-making process. You will watch a part of the interview and fill out a questionnaire. After that, you will watch the second part of the interview, and finally, fill out the last portion of the questionnaire. The process will take approximately 30 to 45 minutes.