TEMPORAL MEMORY ACCURACY FOR AUTOBIOGRAPHICAL EVENTS ACROSS CHILDHOOD

RIYA TRIKHA

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY

YORK UNIVERSITY

TORONTO, ONTARIO

June 2025

Abstract

There is relatively little research about how accurately children remember temporal (when) details associated with autobiographical events. This thesis is part of a large-scale study with a goal of providing legal practitioners information about children's accuracy for several forensically-relevant temporal details (e.g., age during event, month, time of day), and tests whether accuracy is affected by age of the child and valence of the event. Parents nominated and provided timestamped documentation for a positive and negative event (occurring within the past two years). Children (*N*=121; 4-6-years-old, 7-9-year-olds, 10-12-year-olds,13-15-year-olds) then answered various temporal questions about the events. We found age-related improvements for total temporal accuracy (sum of scores across temporal judgments), age, season and month estimates, and higher accuracy for positive compared to negative events for some temporal judgments. These preliminary findings provide novel insight for memory researchers and legal practitioners and will help with understanding time estimates provided by children.

Acknowledgements

I would first like to express my gratitude to my supervisor, Dr. Thanujeni Pathman, for all her support during my graduate school journey and providing me with new opportunities to learn and improve myself. I would also like to thank Tarnpreet Virk and Casey Aurin for their insight, hard work in prepping data material and testing participants with me. I would like to express my gratitude to our collaborators for the project: Drs. Lindsay Malloy, Elaine Reese, Carole Peterson, Qi Wang, Sonja Brubacher, Deidre Brown, J. Zoe Klemfuss, Thomas Lyon and Robert Cribbie. Their expertise and insight are invaluable to every facet of the project. I thank Drs. Lindsay Malloy, Jessica Sutherland and Ayana Thomas for being on my MA defense committee. I would also like to thank my mentors, Drs. Buddhika Bellana and Andrée-Ann Cyr for supporting me and providing me with guidance. I would like to thank the Social Science and Humanities Research Council (SSHRC) for funding this project and providing us with stable support for our long-term initiatives to advancing knowledge in the field.

I would like to thank my family for their unwavering support, encouragement and patience during my graduate school journey; my parents Sanjay and Sonia Trikha, my brother Rohan Trikha and my cousins, Anjali Bhardwaj, Karuna Bhardwaj, Priya Trikha, Sharang Bhardwaj, Gaurav Trikha, Esha Trikha, Reva Trikha and Naiya Trikha. I would also like to thank my friends, William Fisher and Christiane Marie Canillo, for their unconditional kindness and support.

Finally, I would like to thank all parents and caregivers who took the time to participate in the project, and all the wonderful children who worked hard to answer our questions. This project would not be what it is without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Temporal Memory Accuracy for Autobiographical Events Across Childhood

Criminal investigations require legal testimonies from those who have either experienced or witnessed a crime in order to gain more insight into a given case. Of those who testify, many are children as young as 3 years old (e.g., Jones & Krugman, 1986). When children testify, it is important for police interviewers and legal practitioners to use appropriate techniques that produce reliable testimonies and manage expectations about children's recollection of the events. For instance, forensic researchers have built a vast knowledge-base on methods for interviewing children. However, the current knowledge-base is missing information about children's ability to accurately recall temporal details of past events, resulting in little information about how to inquire about and interpret a child's memory recall on these details. Given the importance of time details regarding a legal case, the current study's objective is to add to existing forensic-interviewing knowledge bases by measuring children's temporal memory accuracy, with emphasis on children's ability to place events in time (e.g., judge their age, judge the day, month and season the event took place), for unique past events and track these changes across childhood. Importantly, this project will be using verifiable documentation (e.g., timestamped photographs) as an objective way to verify memory accuracy for events.

Children's Memories and the Justice System

Thousands of children have been, and continue to be, witnesses in various types of legal proceedings, such as criminal, dependency and family. This requires them to give testimonies, which can either build on, or disprove, existing evidence. Their testimonies become especially important in cases in which they are the only person to have directly experienced or witnessed the event in question and can become especially troubling when the crimes are done in secrecy (Bala et al., 2001), such as physical or sexual abuse cases (Goodman et al., 1999). One aspect of past events that witnesses are routinely asked about is the timing of the event (Lyon & Saywitz, 2006), which is paramount for charging specific crimes, creating adequate defenses, and verifying the child's credibility (e.g., State v. Taylor, 2005). However, the general public, as well as legal practitioners, do not put as much emphasis on a child's recount of events, thinking they are not as reliable as those given by adults (Goodman et al., 1984; Yarmey & Jones, 1982, 1983). In fact, details provided in children's testimonies have historically been excluded due to the belief that their distinction between reality and fantasy is distorted (Brainerd & Reyna, 2013; McGough, 1993). Moreover, it is important to acknowledge the plethora of social factors that play a role in a child's eyewitness statement. For instance, research shows that the likelihood of recantations of statements by children increases with factors such as age (i.e., older children are more likely to recant; Gordon et al., 2014; Malloy et al., 2007; Malloy & Mugno., 2016) and familial support (i.e., children with less support are more likely to recant; Malloy et al., 2007; Malloy et al., 2016; Malloy & Mugno., 2016). Furthermore, children are shown to be highly suggestible, especially younger children (Ceci & Bruck, 1993; Lyon et al., 2019). The presence of a social hierarchy during interviews is also shown to affect suggestibility in children. For instance, children's recollections of events may be affected by the types of questions asked (e.g., leading questions; Goodman & Schaaf, 1997) or when they are asked about the event by people in authority positions (e.g., police interviewers; Lampinen & Smith, 1995; Lyon et al., 2019).

Determining the timing of events is important within the forensic context. However, there is uncertainty about children's accuracy when recalling of temporal (time) details for events. Although questions such as "How old were you when [event] happened?" or "When did [event] happen?" are routinely asked by legal practitioners (e.g., Orbach & Lamb, 2007; Powell et al., 2007), police interviewers are unaware at what age children are able to reliably and accurately recall when the event in question occurred. Though legal practitioners do not expect young children to remember the exact date of the event of interest (R. v. B.(G.), 1990), the goal becomes to establish an accurate range of dates in which the event may have occurred (e.g., Richardson, 2006). Indeed, efforts have been made to pinpoint the timing of an event by gauging its relation to a landmark event (e.g., birthdays, holidays; McWilliams et al., 2019) and encouraging the use of temporal landmarks to gain more temporal information during forensic interviews (NICHD; Lamb et al., 2007). However, given there is limited objectively verifiable temporal information for most forensically relevant events, it is difficult to determine how accurate children are when recalling time details for these past events, and how accuracy for temporal information may change across age.

When no estimates are given to forensically relevant questions, literature highlights children's tendency to respond with "I don't know" or to guess. Research shows that the frequency of these types of responses is affected by the format of the question, with children being more likely to respond with "I don't know" to WH questions (e.g., "*who* was there?") than forced-choice questions (e.g., "was your mom with you?"; Andrews et al., 2017; Earheart et al., 2014). However, sometimes, children provide guess responses to questions they do not know the answer to instead of saying "I don't know" (McWilliams et al., 2021). This tendency could reduce the number of accurate responses provided by children and portray their testimonies as unreliable. Given this, some forensic researchers suggest providing children with explicit instructions that encourage the use of "I don't know" responses when appropriate (i.e., ground rules; Lamb et al., 2018). Additionally, young children are shown to possess metacognitive skills to understand and communicate that they do not know (e.g., Rodriguez Steen & Malloy, 2023; Waterman et al., 2000). However, existing literature does not address how the likelihood of children providing “I don’t know” responses and subsequent guess responses is affected by the type of temporal questions asked. Seeing the importance of acquiring information about when an event took place in legal settings, it is essential to know the types of time details children would respond with estimates, “I don’t know” responses and guess responses.

Development of Children’s Temporal Memory

The importance of remembering temporal information extends outside of the legal system, as our memory system as a whole plays an integral role in our daily functioning. Episodic memory is memory for past events that occur in a specific time and place (Tulving, 1972). Autobiographical memory is memory for personally experienced and personally relevant events that occur outside of the lab space (Bauer, 2007). In terms of how episodic and autobiographical memory develop, studies over the years have not only shown that young children can recall specific memories with little to no prompting over extended periods of time (e.g., Fivush et al., 1987; Hamond & Fivush, 1991) but also have shown age-related improvements in recall accuracy. Building on this, developmental scientists have shifted their focus to uncover the types of details children are able to recall within their episodic and autobiographical memories across development.

One type of detail is memory for temporal information, such as temporal order and temporal context, which aids in providing structure and completeness to our memories (Friedman, 2004; Nelson & Fivush, 2004). Interestingly, studies examining children’s memory accuracy for event detail report that memory for temporal details is protracted relative to other types of event details (e.g., Burns et al., 2016; Picard et al., 2012; Scarf et al., 2017). In other words, children display more difficulty recalling time details of an event relative to space (i.e., “where”; Lee et al., 2016) and other contextual details. However, to date, most studies used either objective lab-based stimuli and paradigms (e.g., Scales & Pathman, 2021), staged events (e.g., Friedman & Lyon, 2005), or subjective judgements of parent-reported autobiographical events (e.g., Fivush et al., 1987) when analyzing children's temporal memory accuracy. In other words, there is limited research that gathers objective measures to compare how children remember temporal details of personally experienced past events outside of the lab. Further, studies have mostly examined children's recollection of temporal information using one type of prompt. That is, most studies examining memory for temporal information tend to focus either on children's accuracy for temporal order (e.g., Canada et al., 2020) or temporal context (e.g., McWilliams et al., 2019). In order to gain a comprehensive perspective on children's temporal memory abilities when recalling past events, it is essential to measure the types of temporal information recalled and in multiple ways.

Temporal Order

One type of temporal detail embedded within our memories is temporal order, which involves chronologically ordering (sub-)events in relation to each other (e.g., "X happened before Y"). Remembering the temporal order of events is integral to daily functioning, allowing for skills such as problem-solving, and providing overall coherence to the memories (Habermas & Bluck, 2000). In memory development research, many studies have found that infants are able to successfully order sequences of actions within one event (e.g., Bauer et al., 2000; Carver & Bauer, 1999). However, children are shown to have a harder time ordering events on a wider-scale timeline, such as those spanning over several months (Friedman, 2003). Regardless, most research studies have found that memory performance for temporal order substantially improves throughout childhood (e.g., Friedman, 1991; Lee et al., 2016; Pathman & Ghetti, 2014), albeit slower than other types of memories (e.g., spatial memory; Picard et al., 2012).

Several studies have used lab-based paradigms to analyze temporal order memory. Canada and colleagues (2020) examined temporal order memory in children aged 4-8 years old and found significant age-related improvements, with 8-year-olds reporting significantly better temporal order memory than 4-year-olds. Interestingly, Scales and Pathman (2021) examined participants' temporal order memory using a more naturalistic lab-based stimulus - a story. After participants (4- and 6-year-olds) listened to stories during encoding, experimenters asked forced-choice questions about the occurrence of an event within the story (e.g., "Was it before or after the boy did cartwheels?"). The researchers found temporal order accuracy for both age groups to differ from other proportion scores; memory accuracy for all other spatial and contextual details was higher than that of temporal order for both age groups, replicating literature on the protracted development of temporal order memory. Recently, scientists have been putting more effort into studying the development of temporal order memory in naturalistic settings (i.e., outside of the lab-space). Deker and Pathman (2021) analyzed temporal order memory of children attending a week-long zoo camp, and how temporal order accuracy may be influenced by the temporal distance of a given event. Using the participants' schedules, experimenters administered a primacy task: participants were asked which of the two events they experienced first at the zoo camp that week. The researchers found that only the oldest age group (8- to 10-year-olds) had higher temporal order accuracy for events that were further in distance (long-lag condition). The youngest age groups (4-5- and 6-7-year-olds) did not show any effect of temporal distance on their temporal order accuracy. This study not only replicates the protracted development of temporal order memory but also expands on how the temporal distance of events can impact memory accuracy for when events occur relative to each other.

Though great strides have been made to study the development of temporal order memory, there is relatively less information about how memory for temporal context develops. While acknowledging that the temporal unfolding of different events is forensically relevant, examining how children recall temporal context information, such as how old they were when the event took place, or the month or season that the event took place, is necessary to fill the gap in literature for legal practitioners.

Temporal Context

A critical type of temporal memory is the ability to place events in time or within the appropriate temporal context. For instance, we provide temporal context of an event by associating an event to a specific time point in our lives, such as our age or what grade we were in at the time of the event. Temporal context can also be provided using conventional time scales, such as when we recall the season, month, day or time of day the event happened.

Estimating Age at Time of Event. Since the late 1800s, memory researchers have placed special focus on adults' earliest memories. In fact, studies that have asked participants to judge their age for their earliest memories found strong evidence for a phenomenon later coined as childhood amnesia (Bauer, 2007), in which most adults have trouble recollecting memories from the first 3-4 years of their lives (e.g., Henri & Henri, 1898). Moreover, researchers reliably find another effect called telescoping, in which age judgements for earliest memories are postdated (i.e., participants remember themselves being older than their true age during the event; Huttenlocher et al., 1988; Wang & Peterson, 2014, 2016). Recently, Peterson and colleagues (2024) aimed to explore the question of how early events (ranging from 1 to 10 years ago) are dated by children. Consistent with the telescoping literature, the researchers found that forward telescoping errors were mostly made when children were dating earlier memories. In other words, the earlier the memory, the older the age estimates given by children.

Perhaps more relevant to court cases are studies examining children's abilities to accurately date past events other than earliest memories. Limited studies have asked children to estimate their age for general past events, and studies that do have reported conflicting evidence. For instance, Bauer and colleagues (2007) asked 7- to 10-year-olds to recall a specific, one-time memory from any point in their life related to a cue-word. The researchers found that children were 89% accurate in recalling their age at the time of the events. Wandrey and colleagues (2012) found the contrary; in this forensic study, the experimenters asked 6- to 10-year-old maltreated children to date two specific instances of repeated, emotionally charged events. The researchers found accuracy performance for reported age was around chance, with an average discrepancy score of 87 days. Thus, additional studies are needed to understand when children are able to accurately estimate their age for past events.

Placing Events on a Conventional Time Scale. An essential component of temporal memory and its development is the ability to place events on a conventional time scale. For instance, Friedman (1992) showed that although children as young as 4-years-old are able to identify their birthday month and the current season, their ability to place unique past autobiographical events on a conventional time scale (time of day, day of week, month of year, season) is considerably limited, and is shown to improve from 8 to 9 years old. Moreover, there is conflicting evidence for how accuracy changes for a specific time scale. On the one hand, Friedman and Lyon (2005) had children aged 4- to 13-years-olds experience two staged events and participants were asked to estimate the time the events took place using conventional time scales such as time of day (morning, afternoon, night), month of year and seasons. The researchers found that accuracy for judgements on all time scales improved with age except for the month of the year scale, where they found no age-related improvements in accuracy. On the other hand, Pathman, Larkina and colleagues (2013) examined children's temporal memory by asking them about personal autobiographical events. Children (aged 4-, 6- and 8-years old) were asked to recall four of these autobiographical events selected by the experimenter from the calendar parents populated with unique events over a 4-month period. To collect temporal judgements, experimenters asked children to estimate the time the autobiographical events occurred using four conventional time scales: time of day, day of the week, month of the year and season. The researchers found general age-related improvements for accuracy across all four time scales, but patterns differed for the type of time scale; accuracy improved for day-of-week, month-of-year and season time scales, but no age differences were found for the time-of-day scale.

Given the difference in procedures and paradigms in existing literature, it is difficult to make any definitive conclusions about children's ability to date past events using their age and conventional time scales. Most existing studies that analyze memory accuracy for conventional time scales have used staged events (e.g., Friedman & Lyon, 2005), which does not address temporal memory accuracy for more naturally-occurring, personal events. Although some studies address this concern by using parent- or experimenter-nominated personal events (e.g., Friedman et al., 2011; Pathman, Larkina et al., 2013) and more forensically relevant events (e.g., Wandrey et al., 2012), they have mostly focused on accuracy within a limited age range – young-to-middle childhood. Thus, there is limited information on how memory accuracy for a child's age and conventional time scale details changes across childhood and adolescence, and in comparison to adults.

Factors Impacting Memory: Emotional Valence and Distance of Events

Emotional Valence of Events

A topic of interest in memory research is the effects of emotional valence (positive and negative) of a particular event on memory. In the adult memory literature, several studies using lab-based stimuli and autobiographical memories have found differences in the durability, accessibility and phenomenological richness of valenced memories (e.g., Cahill & McGaugh, 1995; Kensinger & Corkin, 2003). Indeed, literature has consistently found valenced events/stimuli to be remembered with more detail and specificity than neutral events/stimuli (e.g., Ack Baraly et al., 2017; Hamann, 2001), which some argue has an evolutionary basis (rational analysis perspective; Anderson & Schooler, 1991). For instance, the Arousal-Biased Competition theory (Mather & Sutherland, 2011) suggests that although emotion facilitates the binding of high-priority features of the event (e.g., the place, the time, the duration; Mather, 2007; Talmi & Palombo, 2025) by increasing attentional focus, it may impair binding for low-priority stimuli or associations (e.g., context). Interestingly, there is contrary evidence for how recall for positive and negative memories compare. Where some adult studies have found a high number of sensory and other episodic details recalled for both distant positive and negative events compared to neutral events (e.g., Comblain et al., 2005; St. Jacques & Levine, 2007; Wardell et al., 2021), others have found that recall for positive events contain more sensory details than negative and neutral events (e.g., D'Argembeau et al., 2003; D'Argembeau & Van der Linden, 2008, though see Kensinger, 2009). Adding to the nuances of emotionally-valenced memories, recall for these memories has also been found to differ across highly emotional events. For example, research on traumatic (e.g., Porter & Birt, 2001), flashbulb (e.g., Brown & Kulik, 1977; Hirst & Phelps, 2016) and threatening events/memories (e.g., Cliver et al., 2024) show that although those who have experienced these highly negative events produce very detailed and specific recall, the details provided have also been shown to be inaccurate and prone to distortion (e.g., Neisser & Harsch, 1992). Patterns seen in the child memory literature are similar to what researchers have reported about adults' recall of emotionally-valenced memories. For instance, a series of studies found that children's recollections of medical procedures were more detailed and accurate than their memory for more positive or neutral events (e.g., Goodman et al., 1991, 1994; Ornstein, 1995). Moreover, Simpson and colleagues (2023) found that young adolescents recalled self-reported negative events that were relatively unique (i.e., occurred less frequently) were more likely to be recalled than positive events. Similar to adult studies, developmental studies have also found children's recollections of traumatic memories to be less accurate and less coherent than more positive or neutral memories (e.g., Berliner et al., 2003; Bruck et al., 1997; Fivush et al., 2003), perhaps due to children practicing other cognitive avoidance strategies (Koss et al., 1996), resulting in weaker memory traces for the negative events.

Although studies have addressed accuracy of overall details, specificity and coherence of emotional memories (e.g., Berliner et al., 2003), research on the effects of emotional valence on memory for temporal details is small in adult research (see D'Argembeau &Van der Linden, 2005), and untested in developmental research. Given that events discussed within a forensic setting are usually valenced (primarily negative) and temporal landmarks (primarily positive, such as birthdays or recurring holidays; McWilliams et al., 2019) are frequently used to aid in collecting temporal information, it is important to examine how memory for temporal context may differ for positive and negative events.

Temporal Distance of Events

When a certain event occurred relative to the present impacts how well the event is remembered. Both child and adult memory research have consistently reported that more distant or remote events tend to be remembered less vividly (e.g., Campbell et al., 2011; D'Argembeau & Van der Linden, 2004; Johnson et al., 1988) and less coherently (e.g., Van Abbema & Bauer, 2005). This is not surprising; although the rate of forgetting stabilizes over time, older, more remote, events are less often rehearsed and brought to memory than more recent events (e.g., Suengas & Johnson, 1988; Wixted, 2004). For children, researchers have argued that, unlike the power function trend for forgetting older events that is seen in adult literature (e.g., Wixted & Ebbesen, 1991), children's rate of forgetting is different. Indeed, Bauer and colleagues (2007) found that although adults' rate of forgetting slows with the passage of time, children seem to show a continued vulnerability to forgetting. Despite this, some developmental literature shows that children's recall accuracy for distant events is unaffected (e.g., Quas et al., 1999; Van Abbema & Bauer, 2005). Being one of the first to analyze the phenomenon, Hamond and Fivush (1991) asked 37- and 49-month-olds (3- to 4-year-olds) about their Disneyworld trip either 6 months or 18 months after the trip. The researchers found that recall accuracy did not differ across age nor temporal distance from the event, challenging existing memory development findings at the time.

Though these findings were pivotal at the time, they do not reflect the full story. Indeed, research has found that although children's recall is highly accurate in general, the same cannot be said for children's accuracy for temporal information (e.g., Friedman & Lyon, 2005). For instance, when asking children aged 2- to 9-years old about their visit to the emergency room due to a traumatic injury 6 months prior, Peterson and Bell (1996) found that although their recall of the event was accurate, they mentioned very little temporal information in their free recall before the age of 9. Forensically, it is important to acknowledge how memory for events of interest in court cases can be affected by how long ago they occurred. Wandrey and colleagues (2012) asked maltreated 6- to 7-year-olds to recall a specific instance, either their first (which took place on average 2 years ago) or last (which took place on average a year ago) visit to dependency court or their change in placement. The researchers found that regardless of whether the event was their first or last instance, children had a difficult time recalling their age, the month and the season at the time of the given event. However, they found no deficits in children's recall of other details pertaining to the event, further building on the research showing that children's difficulty with temporal memory is not due to their lack of memory for the event.

All in all, there is mixed evidence on how the emotional valence and temporal distance of an event impacts children's memory for those events. For instance, the events nominated for existing autobiographical and forensic studies examining time details provided by children have taken place within a relatively short time frame. This brings into question how accuracy for temporal details may change as past events are more distant. Further, of the events nominated by parents (autobiographical events), a small minority are explicitly negative, making the comparison of temporal accuracy between positively and negatively valenced personal events untested. Given the forensic importance of emotionally-valenced and temporally distant events, it is essential to examine how the temporal distance and emotional valence of unique, personal past events can impact temporal estimates provided by children.

Current Study

This is the first large-scale study to investigate children's (aged 4-15 years old) and adults' memory accuracy for past autobiographical unique events, with a special focus on temporal memory accuracy. Parents were be asked to nominate and describe four separate, unique events: two would have taken place 2-4 years (distant events) in the past and two would have taken place within the past 2 years (recent events). Moreover, each event within this distance manipulation was either positively or negatively valenced. Experimenters also asked parents to provide documentation (e.g., timestamped photographs, text messages) that indicated the time the event took place, allowing the study to have an objective verification of temporal memory accuracy.

Data collection is ongoing. For the purposes of this master's thesis, we focused on the two recent events because the sample size of participants that remembered all events (recent and distant) was relatively low. Moreover, we focused on child groups only because the sample size for the adult group is currently too low. Thus, we examined how age of the child during the session (4-6-year-olds, 7-9-year-olds, 10-12-year-olds, 13-15-year-olds) and emotional valence (positive and negative) of events impact children's accuracy for multiple types of temporal details. Additionally, we examined how the effects of age and valence on accuracy for temporal details compare without participants who responded with "I don't know" and with participants who responded with "I don't know". Subsequent guess responses were not included in this thesis. By examining these factors, we aim to contribute to and build on the existing literature on how memory for temporal details of past events change across childhood. Moreover, we aim to provide relevant information to legal practitioners regarding how to best interpret children's judgement of temporal details mentioned in their testimonies.

Given past literature examining temporal memory accuracy within children, we hypothesize age-related improvements in temporal memory accuracy: as children get older, their judgements for their age at the time of the event will be more accurate, and that there will be general age-related improvements in accuracy for conventional time scale details (season, month, day, time of day) for children 4- to 8-years old (e.g., Pathman, Larkina et al., 2013). Given the mixed findings regarding the effects of emotional valence of events on different types of temporal memory, we have no explicit hypothesis for how accuracy may differ for positive versus negative events. Finally, given the novelty of the project overall, we do not have explicit predictions regarding the interaction between age of child and an event's emotional valence.

Methods

Participants

This thesis sample includes 121 participants: nineteen 4-6-year-olds (*M*=5.16, *SD*=0.83), thirty-two 7-9-year-olds (*M*=7.75, *SD*=0.72), forty-four 10-12-year-olds (*M*=10.82, *SD*=0.79), and twenty-six 13-15-year-olds (*M*=13.96, *SD*=0.96). See Figure B1 (Appendix B) for the age distribution of the final dataset. During data cleaning, experimenters sorted all events as either useable or unusable. Useable events were events that the child successfully recalled and adhered to our criteria (i.e., were unique/one-time events, happened within a 24-hour window, and did not fall on a temporal landmark (e.g., birthday, holiday, etc.)). See Figure B2 (Appendix B) for a detailed summary of the frequency and types of unusable events. Additional participants that were tested but not included in the thesis dataset did not have two useable recent events.

Families were recruited from Canada (21%) and the United States (79%) through several means, such as creating and posting physical and online study-specific recruitment flyers posters in local businesses, social media, email listservs (e.g., Cognitive Development Society listserv), parent groups and other platforms for posting online child studies (e.g., Children Helping Science). As remuneration for completing the study, families received a $30 CAD gift card (Amazon.com, Amazon.ca or Indigo) and were asked if they would like to be entered into a draw for a $200 gift card. Children were also awarded a "Junior Scientist" certificate. Within the current sample, majority of parents who nominated events in session 1 were mothers (93%), had at least one parent in the household with a graduate or professional degree (75%), were White (64%), had a total family income of and above $120,000 CAD (65%). See Figures B3-B6 (Appendix B) for detailed demographic breakdown.

Materials and Procedure

Pre-Screen Consent Form and Event Description Email

All recruitment material (physical and digital posters) contained a link to the study's website (www.childmemorystudy.com), where parents could learn more about the study and the procedures. The "Participate in Study" button on the main study page redirected them to an online survey platform (host: Qualtrics) to complete a pre-screening survey to confirm whether the child falls within the age range of interest. If the child is within the age range, the survey redirected parents/caregivers to the parental consent form. Once parents/caregivers of eligible children consented to participating, they were redirected to schedule the pre-study Zoom session with an experimenter (host: Calendly, Microsoft Bookings). Parents/caregivers also received an email outlining the four events to be discussed during the pre-study survey: a recent positive, distant positive, recent negative and distant negative event. The email outlined definitions for a positive and negative event: A positive event is a memory the child would remember as "good" (i.e., when the child was happy, felt cheerful, excited, or glad), whereas a negative event is a memory the child would remember being as "not so good" (i.e., when the child was unhappy, felt angry, sad, scared, disappointed, or upset). It also emphasized that the events nominated must be unique; the events must have occurred only once. The email also outlined a participant-specific time interval the distant and recent events should fall in, which is determined by the date the parent/caregiver completes the pre-screening survey. A 3-month buffer at each endpoint was included to account for delays in scheduling the second session, which could result in a recent event becoming a distant event or a distant event no longer falling in the time interval. That is, instead of only stating that distant events should have occurred 2-4 years ago, and recent events should have occurred 0-2 years ago, personalized emails gave date ranges such that eligible distant events should have occurred between 27-45 months ago, and eligible recent events 3-21 months ago. For instance, if a parent/caregiver completes the pre-screening survey on October 15th 2024, they were sent an email that instructed them to select distant events that occurred from January 15th 2021 to July 15th 2022 and recent events that occurred from January 15th 2023 to July 15th 2024. Finally, the email stated that parents/caregivers should have these events selected prior to the pre-study session, but experimenters were also able to help parents select appropriate and eligible events during the pre-study session if necessary.

Pre-Study Session: Parent Interview

The pre-study session was a Zoom session with the parent/caregiver and an experimenter. The purpose of this session was to collect information about the four events that the child was directly involved in: a positive recent event, a positive distant event, a negative recent event and a negative distant event. To ensure children are not reminded of any of the events prior to their session (see Main-Study Survey: Child Interview), parents were asked to complete the pre-study session alone (i.e., with their child not in proximity).

Order of event types discussed were randomized. First, parents/caregivers were asked to show documentation (e.g., via the computer camera if the documentation was on another device, via screenshare if the documentation was on their computer) from the event along with a timestamp. For instance, if the event was their child receiving a vaccination, parents/caregivers were able to submit a photograph of an anonymized (i.e., names, date of birth, and other personal information blurred or covered) vaccination card, with the date of the vaccination event nominated. Or, if an event was a family trip to an exhibition, parents/caregivers were able to submit a photograph of the family and/or child at the exhibition, along with the time and date the photograph was taken. Once the documentation was shown and approved as eligible by the experimenter, parents/caregivers were asked to provide their recollection of the event.

Experimenters asked parents several WH-questions about each event; who was at these events, what happened during the events, when and where the events took place, how long the events were (and whether parents have documentation indicating the start and end of the event), how the child felt during the events, and how often they have discussed the events with their child. After describing each event, parents/caregivers were asked to generate a title for the event either using one WHERE detail (e.g., “Canada’s Wonderland” [amusement park]) and one WHAT detail (e.g., “went on the Leviathan” [rollercoaster ride]) or two WHAT details. The generated event title (e.g., “when you went on the Leviathan at Canada’s Wonderland”) would later be used to pinpoint the target event for the child. Additionally, parents/caregivers were asked to generate two reminders or cues: a WHAT detail and a WHO detail (not mentioned in the title). These cues were collected so that experimenters could use them as additional prompts in case the child does not remember the event from the title alone.

After the parent finished providing details for all the events, they were asked to schedule their child’s session. If there was time left over after the session, parents were asked to complete a demographics questionnaire which included questions about their child (age, date of birth, education, ethnicity, languages spoken/understood), the household (income, parental education) and other questions gauging how often the family generally discusses past events. If there was no time left over after the session, parents were emailed the demographics questionnaire to complete, and were asked to complete it preferably before their child's session. The pre-study session took 45-60 minutes to complete.

Main-Study Session: Child Interview

Introduction to Session. The main-study session was a Zoom session with the child and an experimenter. The purpose of this session is to collect information about the child's recollection of the four events nominated by the parents/caregivers. Experimenters acquired verbal assent prior to beginning the session. Next, experimenters went through an understanding check, which lets participants know if they do not understand what the question is asking, that they should respond with "I don't understand", and the experimenter would rephrase the question. To ensure engagement and track progress throughout the session, participants were asked to choose a character (astronaut, rocketship, alien) that went on a trip through outer space that took the same amount of time as the session. The number of planets the character goes to overlapped with the number of tasks participants would be asked to complete (see Figure B7 in Appendix B). Once the character was chosen and the progress tracker was explained, event retrieval began.

Free Recall, Cued Recall and Temporal Memory Questions. Order of event types were presented in the same randomized order as the pre-study session with their parent/caregiver. For each event, experimenters first cued the child with the event title provided by the parent/caregiver in the pre-study session and asked them if they remember this event. If they did not remember the event, children were given an additional WHAT detail, and possibly a WHO detail if the WHAT detail was not sufficient in cueing the event memory. If the event was remembered, experimenters began the set of questions displayed in Figure 1. First, there was a free recall task, during which children were asked to provide an open-ended autobiographical memory recall of the events (i.e., autobiographical narrative) and cued WH-questions (not included in the thesis). After the WH "when" question, participants were asked eight specific temporal memory questions to examine memory for precise temporal details. Children were asked how old they were during the event, what grade they were in during the event and questions homing in on conventional time scale details for the event (i.e., season, month, day, time) and a question about the duration of the event. Following each temporal question, children were asked about their justification/reasoning for their estimates; experimenters asked children how they know the event occurred at that particular time. Then, children were asked how sure they are about their estimates: they rated their confidence about each temporal question on a 3-point visual scale (*not sure*, accompanied by a picture of an animated girl seeming confused; *kinda sure*, accompanied by a picture of the same animated girl appearing to be thinking; *very sure*, accompanied by a picture of the same animated girl smiling with her hands on her hips; Hembacher & Ghetti, 2014). These prompts and questions were presented for each of the four events (Figure 1). Importantly, if children responded with "I don't know" to any temporal question, experimenters asked a follow up question prompting them to "take [their] best guess". Children who responded with "I don't know" for the age estimate were instead asked if the event occurred "before or after [they] were [current age -2]". This thesis only focused on initial responses provided for the temporal questions and guess responses were not analyzed. There was a 3-minute break in between the third and fourth event. During the break, participants were given the option to take either a 3-minute personal break, where they can get a drink of water or use the restroom, or a body break, where the experimenter will put on a 3-minute guided exercise video for the child to follow (https://www.youtube.com/watch?v=-y5yuLuw7KQ).

Figure 1

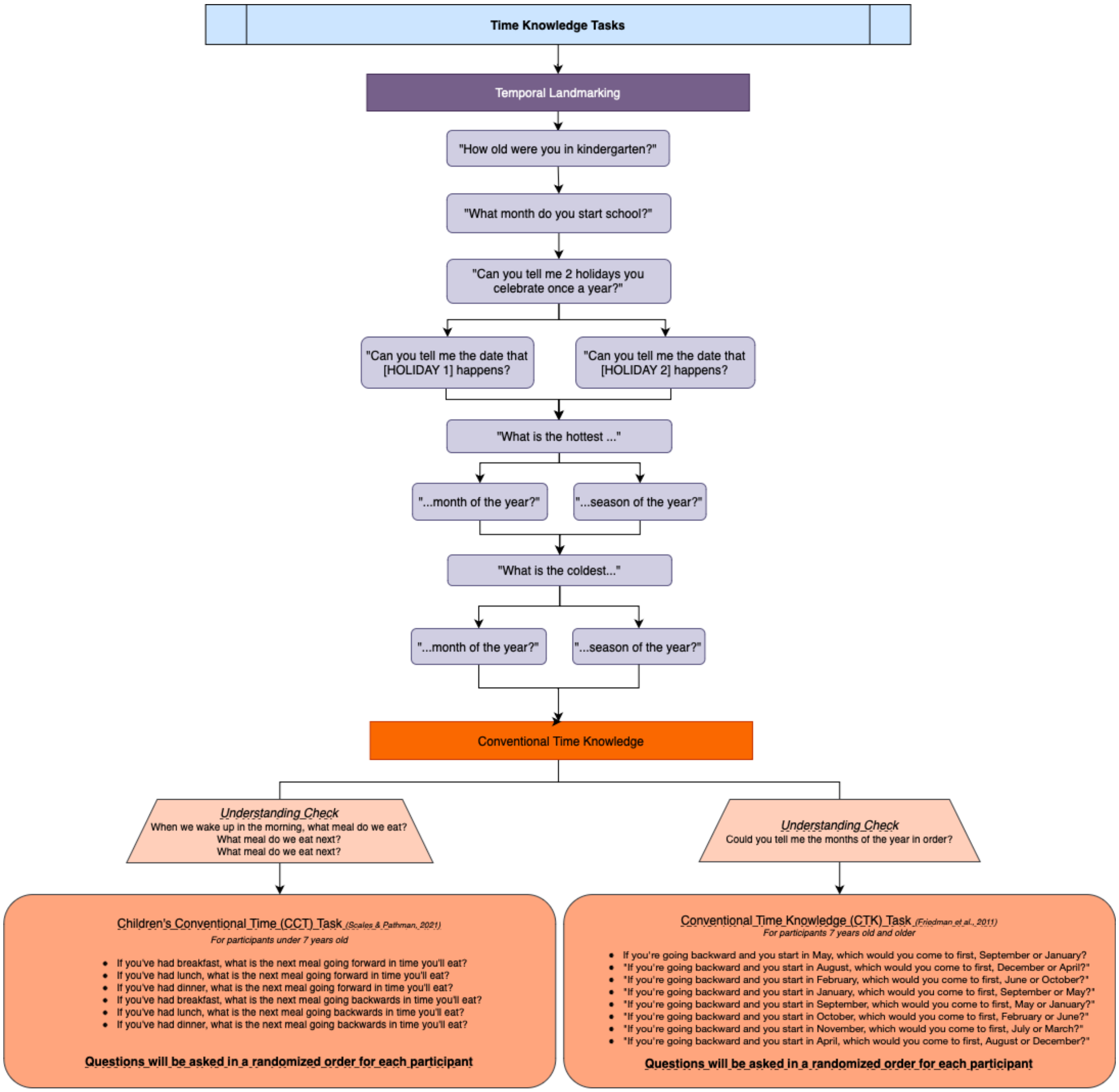
Child Session: Detailed Overview of Event Recall

Note. Current diagram represents questions for one event. All questions were repeated for each unique event. The focus of this thesis is the "*When* did it happen" questions in pink.

Temporal Landmarking Task. After discussing all four events, we measured children's knowledge about time by first assessing their temporal landmarking abilities. See Figure 2 (in purple) for detailed overview of temporal landmarking questions. For the purposes of this thesis, this temporal landmarking task was not analyzed.

Figure 2

Child Session: Detailed Overview of Time Knowledge Tasks



Conventional Time Knowledge Tasks. Next, we assessed conventional time knowledge with either the Children's Conventional Time (CCT) task or the Conventional Time Knowledge (CTK) task. See Figure 2 (in orange) for all questions asked. In order to test conventional time knowledge in children 4- to 6-years old, we used the breakfast subtask from the "flexible retrieval" portion of Scales & Pathman's (2021) CCT task. Participants were asked a total of 6 questions in a randomized order, in which they were required to retrieve which meal of the day (i.e., breakfast, lunch, dinner) would appear first in both a forward and backwards direction.

Three out of the six questions were forward-direction questions (e.g., “if you eat breakfast, what is the next meal you will eat forward in time?”) and the other three were backward-direction questions (e.g., “if you eat dinner, what is the next meal you will eat going backwards in time?”). The proportion of accurate responses (out of six possible) was calculated for each participant. In order to test conventional time knowledge in participants 7 years and older (including adults), we used an adapted version of the “months” task within Friedman’s (1986) larger Conventional Time Knowledge (CTK) task (Pathman et al., 2022). Participants were asked 8 questions in a randomized order, in which they were required to retrieve which month of the year would appear first in a backwards direction. An example question is: “If you’re going backwards and you start in May, which would you come to first, September or January?”. The proportion of accurate responses (out of eight possible) was calculated for each participant.

Scoring, Data Processing and Analysis

Both dichotomous accuracy and deviation scores (i.e., the discrepancy (in units relevant to the question) between the child’s answer and documentation provided; Friedman et al., 2011) were calculated for all temporal questions. Dichotomous accuracy was measured following Pathman, Larkina and colleagues’ (2013) protocol, in which temporal responses were scored as either correct (score of 1) or incorrect (score of 0). Deviation scores were scored following Friedman and colleagues’ (2011) protocol, which measured the absolute, but shortest, discrepancy (in units relevant to the question) between the child’s answer and documentation provided, in which higher deviation scores indicate higher discrepancy from the accurate answer. For instance, if an event occurred in March, but the child recalled it being in January, the child received a dichotomous accuracy score of 0 but received a deviation score of 2. Furthermore, accuracy was compared to chance performance for temporal scales with categorical or finite response options (i.e., season, month, day of the week, time of day) using chi-squared analyses. The expected frequencies for the chi-squared analyses were the numbers of children that would answer correctly and incorrectly if they chose randomly (or equally often) among the options for a given temporal question. For instance, for the seasons of the year (4 possible options), we would expect one out of four children to answer correctly by chance, and three out of four children to answer incorrectly. Only the chi-squared analyses for temporal questions that had an expected frequency higher than 5 (see Howell, 2011) were reported. A detailed outline of how all temporal questions were scored can be found in Table A1 (Appendix A). For the purposes of this master's thesis, justifications and confidence ratings provided for each temporal judgement were not scored.

All the following analyses and visualizations was conducted in RStudio (RStudio Team, 2020). List of functions and packages used for all analyses conducted and visualizations can be found in Table A2 in Appendix A.

Results

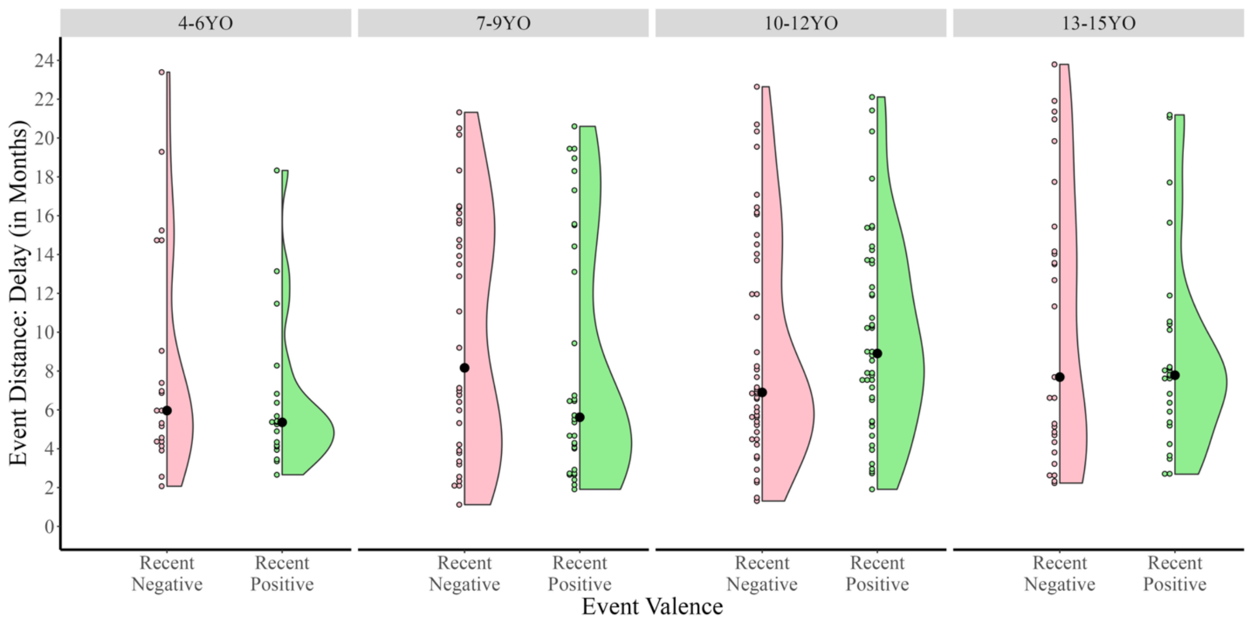
Preliminary Analyses

Delay of Events

The delay of the recent events was calculated as the difference between the date of the recalled event and the date the child session took place. The delays of all memories recalled ranged from 1.12 months to 23.79 months (M=9.08, SD=5.81, Mdn=6.97). See Figure 3 for a visual distribution of delay of events for each age group and valence, and Table A3 (Appendix A) for mean, standard deviation, range and median values of event delay for each age group and event valence. A mixed effects ANOVA analyzing the effects of age group and valence on delay revealed no main effect of age (F(3,234)=1.505, p=0.21), valence (F(1,234)=0.042, p=0.84) nor an interaction, ($F(3,234)=0.776$, $p=0.51$). Thus, any age group or valence differences in temporal memory accuracy cannot be explained by differences in delay.

Figure 3

Delay of Events for Each Age and Event Valence



Note. Black dots in the violin dot plot indicate median values of distribution.

Number of Cues Needed to Recall Events

If the event title was not enough in prompting children to remember the event, experimenters provided two cues, one at a time. Of the useable events, 222 events did not need any cues to be remembered (91%), while 15 events needed at least one cue to be remembered (WHAT-HAPPENED: $n=13$; WHO: $n=2$) and 7 events needed both cues to be remembered. Refer to Table A4 (Appendix A) for a detailed outline of cues used within age group and valence of events.

Temporal Memory Accuracy

Descriptive statistics useful for legal practitioners are provided in Table 1. The table displays the percentage of children that provided accurate responses, inaccurate responses and "I

Table 1

Percentage of Participants with Accurate, Inaccurate and "I Don't Know" Responses Per Age Group, Valence and Temporal Question

		Temporal Questions															
Age Group	**Number of Responses**	*Age estimate*		*Grade estimate*		*Season estimate*		*Month estimate*		*Day estimate*		*Time of day estimate*		*Time on clock estimate*		*Duration estimate*	
		Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive
4-6-year-old (n=19)	Accurate responses	52.63%	63.16%	61.11%	68.42%	35.00%	47.37%	5.00%	15.79%	15.00%	10.53%	55.00%	66.67%	15.00%	5.26%	0.00%	5.26%
	Inaccurate responses	42.11%	26.32%	16.67%	15.79%	55.00%	42.11%	55.00%	36.84%	40.00%	42.11%	40.00%	22.22%	5.00%	31.58%	78.95%	68.42%
	"I don't know" responses	5.26%	5.26%	22.22%	15.79%	10.00%	10.53%	40.00%	47.37%	45.00%	47.37%	5.00%	11.11%	80.00%	63.16%	21.05%	26.32%
7-9-year-old (n=32)	Accurate responses	68.75%	75.00%	68.75%	75.00%	53.13%	62.50%	12.50%	18.75%	28.13%	18.75%	62.50%	62.50%	29.03%	28.13%	3.13%	15.63%
	Inaccurate responses	28.13%	21.88%	25.00%	21.88%	40.63%	31.25%	65.63%	62.50%	53.13%	56.25%	31.25%	37.50%	41.94%	40.63%	90.63%	84.38%
	"I don't know" responses	3.13%	3.13%	6.25%	3.13%	6.25%	6.25%	21.88%	18.75%	18.75%	25.00%	6.25%	0.00%	29.03%	31.25%	6.25%	0.00%
10-12-year-old (n=44)	Accurate responses	70.45%	86.36%	72.73%	81.82%	63.64%	65.91%	34.09%	51.16%	36.36%	31.82%	75.00%	72.73%	38.64%	31.82%	11.36%	13.64%
	Inaccurate responses	25.00%	13.64%	18.18%	13.64%	29.55%	31.82%	56.82%	39.53%	45.45%	59.09%	20.45%	27.27%	43.18%	61.36%	84.09%	84.09%
	"I don't know" responses	4.55%	0.00%	9.09%	4.55%	6.82%	2.27%	9.09%	9.30%	18.18%	9.09%	4.55%	0.00%	18.18%	6.82%	4.55%	2.27%
13-15-year-old (n=26)	Accurate responses	88.46%	96.15%	88.46%	92.00%	77.78%	76.92%	66.67%	57.69%	22.22%	30.77%	70.37%	88.46%	26.92%	42.31%	0.00%	15.38%
	Inaccurate responses	11.54%	3.85%	11.54%	8.00%	18.52%	23.08%	33.33%	34.62%	62.96%	50.00%	29.63%	11.54%	65.38%	53.85%	100.00%	80.77%

"I don't know" responses	0.00%	0.00%	0.00%	0.00%	3.70%	0.00%	0.00%	7.69%	14.81%	19.23%	0.00%	0.00%	7.69%	3.85%	0.00%	3.85%

Note. Total number of participants per age group (denominator for proportion/percentage values) is listed in the "Age Group" column.

This corresponds to the thesis sample size listed in the Methods section don't know" responses for each temporal question. For the oldest age group, there were close to zero "I don't know" responses for all temporal questions except for month of the year, day of the week and time of day estimates. The majority of "I don't know" responses were reported by 4-6-year-olds, and the number of "I don't know" responses decreased as children get older. Given this, analyses for total temporal memory accuracy and dichotomous accuracy of temporal estimates were conducted twice: once excluding children who provided "I don't know" responses and once treating initial "I don't know" responses as not accurate (dichotomous accuracy = 0).

Refer to Table A5 (Appendix A) for the final sample sizes for analyses that excluded and included participants that responded with "I don't know".

Temporal Memory Accuracy Excluding "I Don't Know" Responses

Total Temporal Memory Accuracy. First, accurate dichotomous accuracy judgements for temporal questions were collapsed to create a total temporal accuracy score (see Pathman, Larkina et al., 2013). Total temporal accuracy was reported as the proportion of correctly answered estimates out of the temporal questions answered with an estimate value. For instance, if a participant provided estimates for all eight temporal questions and correctly estimated their age, grade and season during the event, they would receive a score of 3 out of 8 (proportion score=3/8 or 0.375). However, if a participant did not provide an estimate value for two out of the eight temporal questions (i.e., responded with "I don't know", "I don't remember" or did not wish to answer) and correctly estimated their age, grade and season during the event, they would receive a score of 3 out of 6 (proportion score=3/6 or 0.5). See Table 2 for descriptive statistics for total temporal scores per age group and valence and Figure 4 for a boxplot visualization.

A mixed effects ANOVA examining the effects of age and emotional valence on total temporal accuracy found a main effect of age, $F(3,115)=4.29$, $p=0.006$, $\eta^2_p=0.1$, with a Tukey' showing this difference to be significant between the oldest age group (13-15-year-olds) and 4-6-year-olds, $\beta=-0.12$, $t(115)=-2.58$, $p=0.05$, and 7-9-year-olds, $\beta=-0.13$, $t(115)=-3.21$, $p=0.009$. We also found a main effect of valence, $F(1,115)=5.69$, $p=0.01$, $\eta^2_p=0.05$, with a Tukey's showing that positive events had higher total temporal accuracy scores than negative events. No interaction was found between age and valence (Figure 4). A simple linear regression examining the effects of delay on total temporal accuracy showed total temporal accuracy scores to decrease as the delay of events increased, $F(1,1709)=59.62$, $p < 0.001$, $\eta^2=0.03$.

Table 2

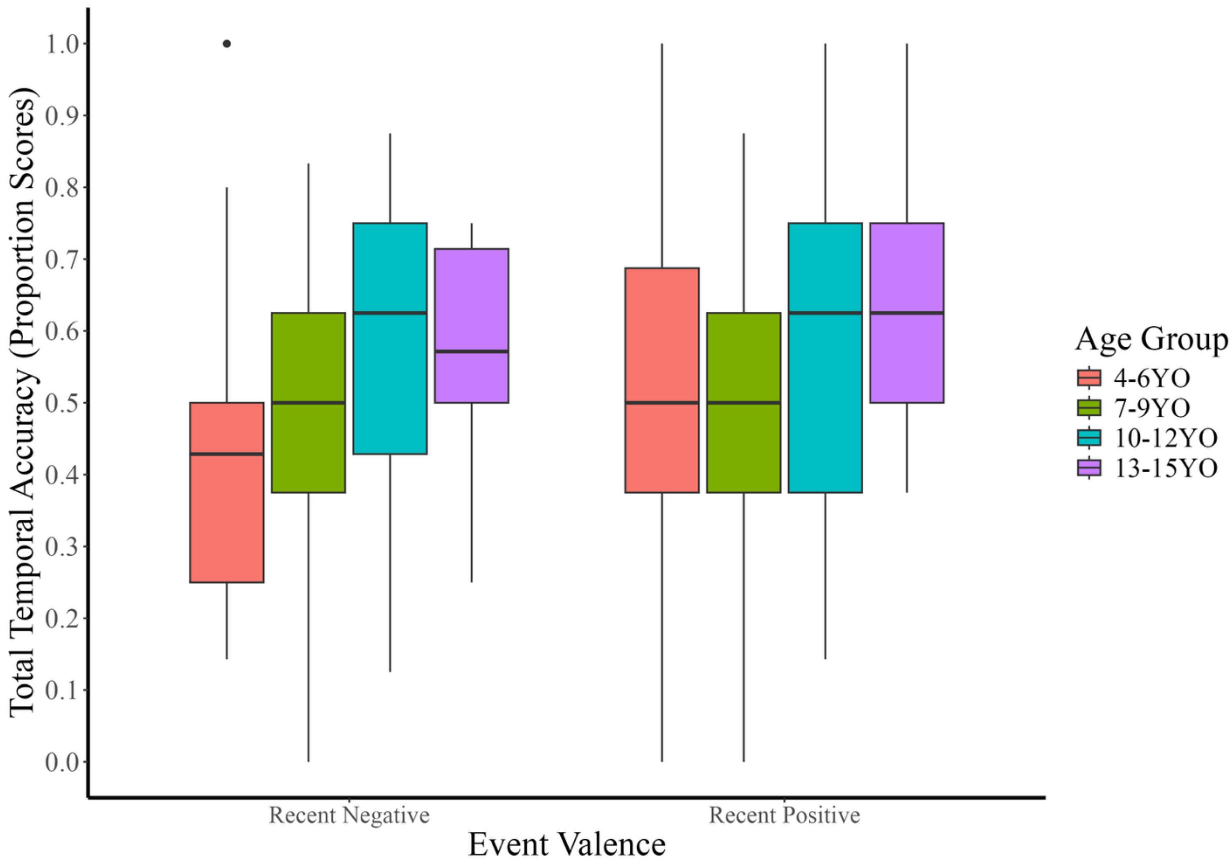
Total Temporal Memory Accuracy Per Age Group and Valence (Excluding "I Don't Know" Responses)

Type of Event (Valence)	Age Group	M	SD	Number of Temporal Questions Included
Recent Negative	4-6-year-old	0.40	0.19	1-8
	7-9-year-old	0.46	0.19	1-8
	10-12-year-old	0.56	0.19	6-8
	13-15-year-old	0.57	0.14	5-8
Recent Positive	4-6-year-old	0.51	0.24	1-8
	7-9-year-old	0.50	0.23	4-8
	10-12-year-old	0.56	0.2	6-8
	13-15-year-old	0.65	0.17	7-8

Note. The values listed are proportion scores of accurate temporal estimate responses out of the total number of temporal questions answered with an estimate. Last column refers to the range of temporal questions answered (i.e., denominator for proportion score). Number of questions included varies because of the exclusion of “I don’t know”, “I don’t remember” and no responses (e.g., “I don’t want to answer”).

Figure 4

Boxplot of Total Temporal Accuracy Scores Within Age and Valence (Excluding “I Don’t Know” Responses)



Note. The horizontal line inside the boxplot represents the median of the distribution. The length of the boxplot represents the interquartile range. The whiskers of the boxplot represent the range of the distribution. The point at 4-6-year-olds’ negative event boxplot is an outlier (observations that are 1.5 times larger than the third quartile of the interquartile range; Wickham, 2016).

Age Estimates. A mixed effects logistic regression examining the effects of age group and valence on dichotomous accuracy for age estimates (refer to Table 3) revealed a main effect of age and valence, but no interaction. Pairwise comparisons (Tukey's) for age revealed that the effect is driven by the difference between the 4-6-year-olds and 13-15-year-olds, showing that, across valence, the odds of having an accurate age estimate increased by 89.3% (OR=0.107; 95% CI = 0.01, 0.82; $p = 0.02$) for 13-15-year-olds compared to 4-6-year-olds. In other words, 13-15-year-olds were significantly more accurate in their age estimates than 4-6-year-olds. Pairwise comparisons (Tukey's) for valence revealed that the odds of having an accurate age estimate increased by 61.3% (OR=0.387; 95% CI = 0.17, 0.88; $p = 0.02$) for positive events compared to negative events. In other words, across age, age estimates for positive events were significantly more accurate than age estimates for negative events. A logistic regression examining the effects of delay on accuracy was significant, showing that the odds of having an accurate age estimate decreased by 9.63% (OR=0.9; 95% CI = -0.15, -0.05; p < 0.001) for every unit increase of delay. In other words, accuracy for age estimates decreased as the delay of events increased.

A mixed effects ANOVA to examine the effects of age group and valence on deviation accuracy (refer to Table 3) revealed a main effect of age, $F(3,112)=4.21$, $p=0.007$, $\eta^2_p=0.1$, and valence, $F(1,112)=8.4$, $p=0.004$, $\eta^2_p=0.07$, but no interaction. Pairwise comparisons (Tukey's) for age revealed that 4-6-year-olds had significantly higher deviation scores than 13-15-year-olds, $\beta=4.26$, $t(112)=3.37$, $p=0.006$. Pairwise comparisons (Tukey's) for valence revealed that deviation scores were significantly higher for negative events compared to positive events, $\beta=1.92$, $t(112)=2.9$, $p=0.006$. A simple linear regression examining the effects of delay on accuracy showed deviation scores increased as the delay of events increased, $F(1,236)=16.19$, $p < 0.001$, $\eta^2=0.06$.

Table 3

Dichotomous and Absolute Deviation Accuracy Per Age Group, Valence and Temporal Question.

Age Group	Accuracy Scale	Temporal Questions															
		Age estimate		*Grade estimate*		*Season estimate*		*Month estimate*		*Day estimate*		*Time of day estimate*		*Time on clock estimate*		*Duration estimate*	
		Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive
4-6-year-old (n=19)	Dichotomous accuracy (%)	53%	70%	78%	81%	42%	53%	8%	30%	27%	20%	58%	75%	25%	14%	0%	9%
	Absolute deviation scores	6.32 (7.34)	3.33 (5.53)	0.43 (1.09)	0.19 (0.4)	0.63 (0.6)	0.71 (0.85)	2.67 (2.02)	2.6 (2.6)	1.18 (0.98)	1.3 (1.06)	0.47 (0.61)	0.25 (0.45)	110.5 (76.69)	351.43 (235.88)	232.75 (254.9)	216.82 (232.95)
	Number of responses (*n*)	19	17	16	16	18	17	12	10	11	10	19	16	4	7	18	13
7-9-year-old (n=32)	Dichotomous accuracy (%)	71%	77%	73%	77%	57%	70%	16%	23%	35%	25%	67%	62%	41%	41%	3%	16%
	Absolute deviation scores	4.26 (7.3)	2.71(5.1)	0.3 (0.53)	0.23 (0.43)	0.5 (0.63)	0.39 (0.62)	2.16 (1.75)	1.35 (1.16)	1.42 (1.24)	1.54 (1.18)	0.4 (0.62)	0.44 (0.62)	131.68 (191.6)	142.36 (205.94)	163.97 (284.4)	103.5 (122.3)
	Number of responses (*n*)	31	31	31	32	30	30	25	26	26	24	30	32	23	22	30	32
10-12-year-old (n=44)	Dichotomous accuracy (%)	74%	86%	80%	86%	69%	66%	37%	56%	44%	35%	78%	73%	53%	34%	12%	14%
	Absolute deviation scores	3.14 (5.34)	1.64 (4.17)	0.22 (0.48)	0.17 (0.44)	0.38 (0.62)	0.45 (0.7)	1.3 (1.45)	0.95 (1.54)	1.03 (1.16)	1.18 (1.11)	0.29 (0.64)	0.32 (0.56)	80.5 (155.23)	127.85 (191.31)	177.15 (272.33)	111.84 (170.26)
	Number of responses (*n*)	42	44	40	42	41	43	40	40	36	40	42	44	36	41	42	43
13-15-year-old (n=26)	Dichotomous accuracy (%)	85%	96%	88%	92%	77%	77%	67%	62%	26%	38%	70%	88%	29%	44%	0%	16%
	Absolute deviation scores	1.78 (4.34)	0.46 (2.35)	0.12 (0.33)	0.08 (0.28)	0.26 (0.53)	0.27 (0.53)	0.78 (1.55)	1.08 (1.91)	1.22 (0.95)	0.9 (0.83)	0.37 (0.63)	0.12 (0.33)	136 (160.34)	59.24 (100.73)	309.9 (639.13)	72.81 (83.81)
	Number of responses (*n*)	26	26	26	26	26	26	26	24	23	21	26	26	25	25	25	26

Note. Absolute deviation scores represent the absolute difference between the participant's estimate and the correct answer and only obtained for non-cyclical temporal questions/scales. Mean deviation scores are in the units of the temporal question: in months for age estimates, in grades for grade estimates, in seasons for season estimates, in days for day estimates, in time-of-day estimates, in minutes for time-on-clock estimates, in minutes for duration estimates. Standard deviations are reported in parentheses.

Grade Estimates. A mixed effects logistic regression examining the effects of age group and valence on dichotomous accuracy for grade estimates (refer to Table 3) revealed no significant main effects or an interaction. Thus, we did not find evidence indicating that grade estimates differed across age or valence. A logistic regression examining the effects of delay on dichotomous accuracy was significant, showing that the odds of having an accurate grade estimate decreased by 7.28% (OR=0.93; 95% CI = -0.13, -0.02; $p < 0.001$) for every unit increase of delay. In other words, accuracy for grade estimates decreased as the delay of events increased.

A mixed effects ANOVA examining the effects of age group and valence on deviation accuracy (refer to Table 3) revealed no main effects of age or valence, nor an interaction. A simple linear regression examining the effects of delay on accuracy was only marginally significant, showing trends of deviation scores being higher as the delay for events increased, $F(1,222)=3.77$, $p=0.053$, $\eta^2 = 0.02$.

Season Estimates. A mixed effects logistic regression examining the effects of age group and valence on dichotomous accuracy for season estimates (refer to Table 3) revealed a main effect of age, but no main effect of valence nor an interaction. Pairwise comparisons (Tukey's) for age revealed that the effect is driven by the difference between the 4-6-year-olds and 13-15-year-olds, showing that, while holding valence constant, the odds of having an accurate season estimate increased by 73.5% (OR=0.265; 95% CI = 0.08, 0.88; $p = 0.02$) for 13-15-year-olds compared to 4-6-year-olds. In other words, 13-15-year-olds were significantly more accurate in their season estimates than 4-6-year-olds. A logistic regression examining the effects of delay on season accuracy was not significant, indicating that accuracy for season estimates did not change or differ across delay of events.

A mixed effects ANOVA examining the effects of age group and valence on deviation accuracy (refer to Table 3) revealed a main effect of age, $F(3,111)=3.19$, $p=0.05$, $\eta^2_p = 0.07$. Pairwise comparisons (Tukey's) for age revealed that the effect is driven by the difference between the 4-6-year-olds and 13-15-year-olds, showing that deviation scores were significantly higher for 4-6-year-olds compared to 13-15-year-olds, $\beta=0.39$, $t(111)=2.8$, $p=0.03$. A simple linear regression examining the effects of delay on accuracy was not significant.

Chi-squared analyses examining chance performance revealed that 7-9-year-olds, 10-12-year-olds and 13-15-year-olds recalled the season correctly more often than would be expected by chance for both negative and positive events: negative events, X^2 (1) = 16.04, $p < 0.001$; X^2 (1) = 40.98, $p < 0.001$; X^2 (1) = 43.13, $p < 0.001$, for 7-9-year-olds, 10-12-year-olds and 13-15-year-olds respectively; positive events, X^2 (1) = 27.78, $p < 0.001$; X^2 (1) = 41.31, $p < 0.001$; X^2 (1) = 37.38, $p < 0.001$, for 7-9-year-olds, 10-12-year-olds and 13-15-year-olds respectively. Chi-squared analyses for 4-6-year-olds were not computed because expected frequency was below 5 given sample size.

Month Estimates. A mixed effects logistic regression examining the effects of age group and valence on dichotomous accuracy for month estimates (refer to Table 3) revealed a main effect of age. Pairwise comparisons (Tukey's) for age revealed that the effect is driven by the difference between the two youngest groups (4-6-year-olds and 7-9-year-olds) and the 13-15-year-olds. More specifically, while holding valence constant, the odds of having an accurate month estimate increased by 91% (OR=0.09; 95% CI = 0.01, 0.67; $p = 0.01$) for 13-15-year-olds compared to 4-6-year-olds and by 89% (OR=0.11, 95% CI = 0.03, 0.46; $p < 0.001$) for 13-15-year-olds compared to 7-9-year-olds. In other words, 13-15-year-olds were significantly more accurate in their month estimates than both 4-6-year-olds and 7-9-year-olds. Next, a logistic regression examining the effects of delay on month accuracy was not significant.

A mixed effects ANOVA examining the effects of age group and valence on deviation accuracy (refer to Table 3) revealed a main effect of age, $F(3,87)=4.67$, $p=0.004$, $\eta^2_p = 0.14$. Pairwise comparisons (Tukey's) for age revealed that the effect is driven by the difference between the 4-6-year-olds and the two oldest age groups (10-12-year-olds and 13-15-year-olds), showing that deviation scores were significantly higher for 4-6-year-olds compared to 10-12-year-olds, $\beta=1.46$, $t(87)=3.01$, $p=0.02$, and 13-15-year-olds, $\beta=1.65$, $t(87)=3.24$, $p=0.009$. A simple linear regression examining the effects of delay on accuracy was not significant.

Chi-squared analyses examining chance performance for month estimates were not computed because expected frequency for all age groups was below 5.

Day of Week Estimates. A mixed effects logistic regression examining the effects of age group and valence on dichotomous accuracy for day of the week estimates (refer to Table 3) revealed no main effects or interactions. Moreover, a logistic regression examining the effects of delay on day of the week accuracy was not significant.

A mixed effects ANOVA examining the effects of age group and valence on deviation accuracy (refer to Table 3) revealed no main effects or interactions. A simple linear regression examining the effects of delay on accuracy was not significant.

Chi-squared analyses examining chance performance revealed that 10-12-year-olds recalled the day of the week correctly more often than would be expected by chance for both negative and positive events: negative events, X^2 (1) = 26.74, $p < 0.001$; positive events, X^2 (1) = 14.01, $p < 0.001$. Chi-squared analyses for 4-6-year-olds, 7-9-year-olds and 13-15-year-olds were not computed because expected frequency was below 5.

Time-of-day Estimates. A mixed effects logistic regression examining the effects of age group and valence on dichotomous accuracy for time-of-day estimates (refer to Table 3), which revealed no main effects or interactions. Moreover, a logistic regression examining the effects of delay on time-of-day accuracy was not significant.

A mixed effects ANOVA examining the effects of age group and valence on deviation accuracy (refer to Table 3) revealed no main effects or interactions. A simple linear regression examining the effects of delay on accuracy was not significant.

Chi-squared analyses examining chance performance revealed that 7-9-year-olds, 10-12-year-olds and 13-15-year-olds recalled the time of day correctly more often than would be expected by chance for both negative and positive events: negative events, X^2 (1) = 27.78, $p <$ 0.001; X^2 (1) = 64.28, $p < 0.001$; X^2 (1) = 29.64, $p < 0.001$, for 7-9-year-olds, 10-12-year-olds and 13-15-year-olds respectively; positive events, X^2 (1) = 24, $p < 0.001$; X^2 (1) = 53.45, $p <$ 0.001; X^2 (1) = 55.84, $p < 0.001$, for 7-9-year-olds, 10-12-year-olds and 13-15-year-olds respectively. Chi-squared analyses for 4-6-year-olds were not computed because expected frequency was below 5.

Time on the Clock Estimates. A mixed effects logistic regression examining the effects of age group and valence on dichotomous accuracy for time estimates (refer to Table 3) revealed no main effects or interactions. Moreover, a logistic regression examining the effects of delay on time accuracy was not significant.

A mixed effects ANOVA examining the effects of age group and valence on deviation accuracy (refer to Table 3) revealed no main effects nor an interaction. A simple linear regression examining the effects of delay on accuracy was not significant.

Duration Estimates. Before discussing the results, it is important to note that 2 duration estimates (one from a 7-9-year-old, one from a 13-15-year-old, both estimates were for negative events) were removed from the dataset, as they were more than 3 standard deviations away from the mean duration estimates (*M*=477.24, *SD*=3029.06). These duration estimates were 43,800 minutes and 10,800 minutes.

A mixed effects logistic regression examining the effects of age group and valence on dichotomous accuracy for duration estimates (refer to Table 3) revealed no main effects nor an interaction. Moreover, a logistic regression examining the effects of delay on duration accuracy was not significant.

A mixed effects ANOVA examining the effects of age group and valence on deviation accuracy (refer to Table 3) revealed a main effect of valence, $F(1,102)=4.18$, $p=0.04$, $\eta^2_p = 0.04$. Pairwise comparisons (Tukey's) for valence revealed that deviation scores for duration estimates were significantly longer for negative events than positive events, $\beta=87.1$, $t(102)=2.04$, $p=0.04$. A simple linear regression examining the effects of delay on accuracy was not significant.

Exploratory Analyses: Directionality of Age, Grade and Duration Estimates. In addition to calculating deviation scores as absolute differences, we also explored directionality. Using the scoring protocol in Peterson and colleagues (2024), we calculated these directional deviation scores by taking the temporal estimate provided by a child for each event and subtracted the correct temporal estimate. So, positive deviation scores would mean that children's estimates were larger than the correct estimate (overestimations), and negative deviation scores would mean that children's estimates were smaller than the correct estimate (underestimations). Seeing that directionality of temporal estimations can only be calculated for non-cyclical temporal scales, we conducted mixed effects ANOVAs examining the effects of age and emotional valence on directional deviation scores for age estimates, grade estimates and duration estimates. The mean directional deviations for each age group across valence of events can be found in Table A6 (Appendix A).

The mixed effects ANOVAs measuring the effect of age group and valence on directional deviation scores for age, grade and duration estimates revealed no main effects or interaction. Thus, we did not find any evidence that directional deviation scores differenced by age or valence.

Temporal Memory Accuracy Including "I Don't Know" Responses

Analyses involving dichotomous accuracy were rerun after including "I don't know" responses scored as zero. This allowed us to examine accuracy performance for the full sample of thesis participants and included all eight temporal questions. For most of the analyses, results replicated the analyses conducted when excluding "I don't know" responses, with the exception of total temporal accuracy and time on clock estimates as described below.

Descriptive statistics for total temporal memory scores can be found in Table 4. A mixed effects ANOVA examining the effects of age group and valence on total temporal accuracy found the main effect of age group to be stronger, $F(1,115)=15.12$, $p<0.001$, $\eta^2_p = 0.31$, with a Tukey's showing this difference to be significant between all age groups except for between the two oldest age groups (10-12 year olds and 13-15 year olds). In other words, when we included all responses and participants, we now see that 10-12-year-olds and 13-15-year-olds reported higher total temporal memory scores than 4-6-year-olds (β=-0.23, $t(109)$=-5.46, $p<0.001$; β=-0.29, $t(109)$=-6.23, $p<0.001$, for 10-12-year-olds and 13-15-year-olds respectively) and 7-9-year-olds (β=-0.1, $t(109)$=-2.97, p=0.02; β=-0.16, $t(109)$=-4.06, $p<0.001$, for 10-12-year-olds and 13-15-year-olds respectively). Moreover, 7-9-year-olds reported higher total temporal memory scores than 4-6-year-olds (β=-0.13, $t(109)$=-2.89, p=0.02). There was no significant difference in total temporal memory scores between 10-12-year-olds and 13-15-year-olds.

A mixed effects logistic regression examining the effects of both age group and valence on time on clock estimates revealed a main effect of age, with a Tukey's showing that, while holding valence constant, the odds of having an accurate time estimate increased by 85% (OR=0.15; 95% CI = 0.04, 3.01; p = 0.03) for 10-12-year-olds compared to 4-6-year-olds. In other words, when including all responses and participants, 10-12-year-olds were significantly more accurate in their time on clock estimates than 4-6-year-olds.

Table 4

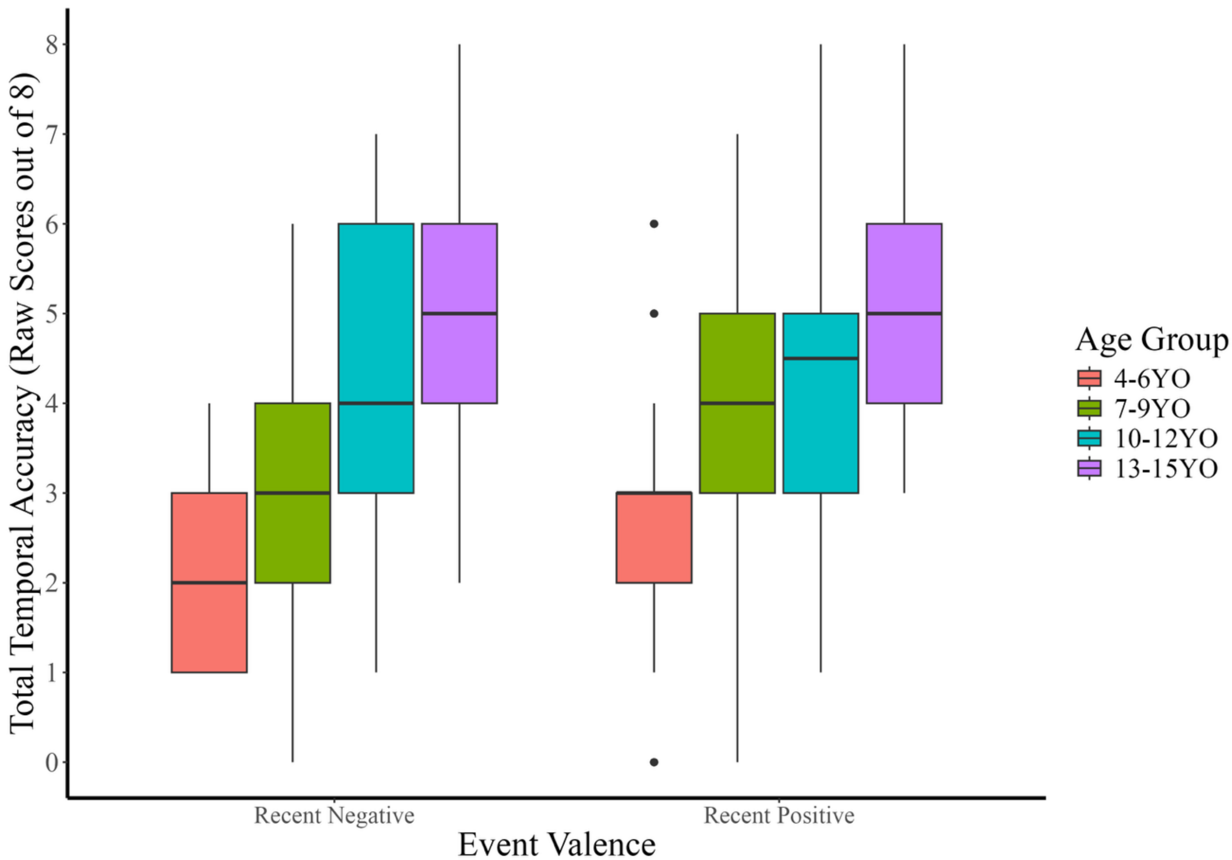
Total Temporal Score Per Age Group and Valence (Including "I Don't Know" Responses as Not Correct)

Type of Event (Valence)	Age Group	M	SD	Range
Recent Negative	4-6-year-old	2.27	1.13	1-4
	7-9-year-old	3.17	1.44	0-6
	10-12-year-old	4.24	1.58	1-7
	13-15-year-old	4.73	1.46	2-8
Recent Positive	4-6-year-old	2.89	1.33	0-6
	7-9-year-old	3.81	1.78	0-7
	10-12-year-old	4.38	1.56	1-8
	13-15-year-old	4.92	1.32	3-8

Note. Range column indicates range of total temporal memory scores in the sample. Maximum possible score is an 8 for all groups.

Figure 5

Boxplot Of Total Temporal Accuracy Scores Within Age and Valence (Including "I Don't Know" Responses as Incorrect)



Note. The horizontal line inside the boxplot represents the median of the distribution. The length of the boxplot represents the interquartile range. The whiskers of the boxplot represent the range of the distribution. The points at 4-6-year-olds' positive event boxplot are outliers (observations that are 1.5 times larger than the third quartile of the interquartile range; Wickham, 2016).

Conventional Time Knowledge Performance

The average CCT task performance for the 4-6-year-olds and average CTK task performance for the older age groups can be found in Table 5.

A one-way ANOVA examining the effect of age group on CTK performance found a significant effect of age, $F(2,92)=7.28$, $p=0.001$, $\eta^2=0.14$. Pairwise comparisons (Tukey's) revealed that the effect is driven by the difference between the 7-9-year-olds and the two oldest age groups, showing that CTK scores were significantly higher for 10-12-year-olds, β=0.19, t(93)=3.16, p=0.006, and 13-15-year-olds, β=0.24, t(92)=3.43, p=0.002, compared to 7-9-year-olds. The two oldest groups did not differ in their CTK performance.

Table 5

Conventional Time Knowledge Task Performance Across Age

Age Group	Time Knowledge Task	M	SD	Range
4-6-year-old	CCT (max score=6)	3.22 (0.54)	1.66 (0.28)	0, 5 (0, 0.83)
7-9-year-old	CTK (max score=8)	4.9 (0.61)	1.81 (0.23)	2, 8 (0.25, 1)
10-12-year-old	CTK (max score=8)	6.43 (0.8)	2.32 (0.29)	0, 8 (0, 1)
13-15-year-old	CTK (max score=8)	6.83 (0.85)	1.67 (0.21)	2, 8 (0.25, 1)

Note. Mean, standard deviation and range values display the value for the task, followed by the proportion values in the parentheses. Proportion values are listed to aid in interpreting task performance across age/different time knowledge task.

Discussion

The purpose of the current study was to investigate temporal memory accuracy for recent (occurred within the past two years) autobiographical events. We examined how accuracy is impacted by age of the child by comparing four age groups (4-6-year-olds, 7-9-year-olds, 10-12-year-olds, 13-15-year-olds) and testing the effects of emotional valence (negative and positive) of events. Experimenters asked eight temporal questions about the event, which required children to estimate their age and grade at the time of the event and make conventional time and duration judgements about the event. This study provides novel insight into temporal memory accuracy across childhood because as opposed to staged or lab-based events, children were asked about unique, naturally occurring autobiographical events. Additionally, instead of solely relying on parents'/caregivers' temporal judgement for past events, we verified the date and time of the events using timestamped documentation, allowing for more objective measures of accuracy. In line with our hypothesis, we found age-related improvements for some temporal memory questions. Interestingly, we also found valence differences for some temporal memory questions. Though most of our findings did not differ between analyses excluding and including participants that responded with "I don't know", there were some additional main effects found once participants who responded with "I don't know" were included back in the dataset.

Most interesting to legal practitioners is the findings that the number of "I don't know" responses differed for each temporal question, as well as for each age group and event valence. For instance, the least amount of initial "I don't know" responses were reported for age estimates. This may suggest that children feel more confident in providing an estimate for how old they were during a past event (irrespective of accuracy) because it is a time scale they are the most familiar with. On the other hand, there were increased "I don't know" responses for temporal scales that are taught in school curriculums (e.g., months of the year, days of the week, time on clock), and were mostly reported by 4-6-year-olds. This may simply suggest that young children may not yet have the knowledge to produce estimates for these time scales. Legal practitioners may use these findings to modify which temporal questions are asked to which age groups, as well as manage expectations for the types of responses they may receive for certain temporal questions. For instance, if our results for children hold after data collection is complete, our findings suggest that legal practitioners are encouraged to inquire about time details that are embedded into rich contextual details, such as their age and what grade/school year they were in during the event, as children across all age groups performed well above chance for these temporal questions. However, our findings also suggest that, in efforts to reduce “I don’t know” responses from children, legal practitioners should avoid temporal questions that require more formal education to know, such as the time on the clock during the event.

Our first set of preliminary results showed that the distribution of event delay is positively skewed, in that most of the events across age and valence occurred less than a year ago, with median values for event delay for all age groups and valence being below 10 months (see Table A3 in Appendix A). Importantly, delay did not vary by age nor event valence. More forensically relevant findings are the second set of preliminary results, showing that most children were able to recall an event by prompting them with the title of the event alone (which included either a WHAT-HAPPENED and WHERE detail or two WHAT-HAPPENED details). This suggests that even children as young as 4-years-old may not need multiple cues in order to remember an event occurring within the past two years, corroborating seminal research on children’s abilities to recall specific memories with little to no prompting (e.g., Fivush et al., 1987; Hamond & Fivush, 1991).

We also found accuracy for total temporal accuracy scores, age estimates, season estimates, and month estimates to differ across age. These findings complement findings from Pathman, Larkina and colleagues’ (2013), who found age-related improvements in both total temporal accuracy in 4-, 6-, and 8-years old, as well as individual temporal questions regarding season. Interestingly, unlike Pathman, Larkina and colleagues (2013), we found all age groups to recall time of day estimates accurately more often than what would be expected by chance for both negative and positive events. This builds on previous literature on temporal memory abilities by showing that children as young as 4-years old are able to recall even precise time scales, like time of day, well above what would be expected by chance.

For age estimates in particular, our findings not only complement those of Bauer and colleagues (2007), who also found children aged 7- to 10-years old to be highly accurate (89% accuracy rate) when recalling their age for a specific, one-time memory, but also extend them. Indeed, our findings show that children as young as 4-years-old are able to correctly recall their age at the time of an event more than half of the time (61.5% dichotomous accuracy rate across valence). Our measure of absolute deviation accuracy supports this finding, showing that even when children were not *exactly* correct (dichotomous accuracy $\neq 1$) at estimating their age, they were only 2.95 months (across age group and valence), on average, off from the correct age estimate. Together, these findings suggests that children as young as 4-years-old, for the most part, are able to correctly estimate how old they were for past events that have occurred within the past two years, and accuracy for age estimates only increases as children get older.

The lack of significant findings for grade estimates may be attributed to the ceiling effect in accuracy scores; our findings show that children across all age groups and valence accurately remembered what grade they were in for over 80% of the events. High accuracy for grade estimates may be explained by the increased contextual information attached to a given school-grade relative to other types of temporal information, such as their teacher during the school year, their classrooms and the subjects they learned about (Friedman, 1991). Seeing that context plays an important role in both facilitating encoding of incoming information and aiding in retrieving certain information by acting as a cue (Edgin et al., 2015; Hayes, Nadel & Ryan, 2007), remembering what grade children were in during a past event may be easier than trying to recall information that lacks additional contextual information. This speculation may be explained further once we examine the justifications provided by children for their grade estimates. Interestingly, the lack of significant findings for duration estimates may be attributed to the floor effect in accuracy scores; our findings show that children across all age groups and valence accurately remembered the duration for only 9% of the events. Several research studies investigating development of duration judgements have found young children to overestimate duration of events, attributing it to factors such as underdeveloped knowledge for duration units (Pouthas, 1993), whereas older children and adults tend to underestimate duration of events (temporal compression; Jeunehomme & D’Argembeau, 2018). However, Block and colleagues (1999) showed that teenagers (> 15-years-old) performed similarly to adults in their duration estimates, supporting that accuracy for duration estimates becomes adult-like closer to adolescence.

Time on clock estimates did not significantly differ between age groups, but effects for dichotomous accuracy were seen once “I don’t know” responses were included. Specifically, younger children were less correct for their time on clock estimates than older children. Inclusion of “I don’t know” responses introduced more responses that were “not correct” as opposed to inaccurate, suggesting that 4-6-year-olds were more uncertain when producing a time on clock estimate than older children for past events. This may be explained by the underdeveloped knowledge of the clock system up until fourth and fifth grade, or 10-12-years old (Friedman & Laylock, 1989).

We observed effects for total temporal accuracy scores, age estimates and duration estimates for valence of events, with estimates being more accurate for positive than negative events. This suggests that children across all ages had a difficult time estimating these temporal details for negative events. Our findings complement research on the effects of distortion of time perception, namely duration, for negative stimuli (e.g., Droit-Volet & Meck, 2007; Gil & Droit-Volet, 2011; Zakay & Block, 1997). Wandrey and colleagues (2012) had similar results for their negative events (placements and court visits) in children aged 6- to 7-years old, showing that only half of their sample (n=167) was able to correctly recall their age at the time of their visit, with their estimates being off by 6 to 12 months. Though the absolute deviation accuracy reported by Wandrey and colleagues is higher than what we reported (negative events: 3.87 months), both findings suggest that children may find it more difficult to report their age at the time of negative events compared to positive events. Further, literature on general recall of valenced events may support our findings. For instance, Fivush and colleagues (1993) showed that children's (5-12-years-old) recall for positive events was more descriptive and contained more details about objects and people than negative events, which contained more emotional and evaluative details. Seeing that age estimates are descriptive details about a person, it may be that children tend to better remember these details for positive events than for negative events. Another explanation may be that negative events are more difficult to recall due to weaker memory traces as a consequence of cognitive avoidance strategies utilized by children (Koss et al., 1996). The effect may also be driven by the frequency in which families discuss negatively-valenced events. Indeed, research examining memory rehearsal have shown that people tend to discuss or reminisce about positive memories more often than negative ones (e.g., Walker et al., 2009), resulting in more vivid and detailed memory recollections for highly-rehearsed events (e.g., Svoboda & Levine, 2009). Similarly, studies have shown children who have parents (primarily mothers) that reminisce with them about past events in an elaborative way to have better recall for those events (e.g., Hudson, 1993; Reese, Haden & Fivush, 1993). We did not find a significant effect of valence for the remaining temporal questions, which is reflected by similar accuracy scores for negative and positive events across age groups.

Moreover, we found an effect of delay of an event for age estimates and grade estimates, showing that estimates were less accurate for events that occurred further in time. This is consistent with previous literature showing that memory performance deteriorates with time (e.g., Friedman et al., 2011; Van Abbema & Bauer, 2005). Though, the lack of significant findings for the remaining temporal questions, which was also found by Pathman, Larkina and colleagues (2013), may support previous findings that show children to be similarly detailed in their recall of distant events as recent events (Bauer et al., 2007; Fivush et al., 1987). However, it is important to note that the range of delay in the current dataset was narrow, so it may be possible for effects of delay to be present once we include more distant events (i.e., those that took place 2-4-years ago).

Limitations and Future Directions

Although the findings for the current dataset already provide novel insight, there is more work to be done in order to produce more reliable findings. First, our sample size is unbalanced, with a smaller number of 4-6-year-olds and 13-15-year-olds than the 7-9-year-olds and 10-12-year-olds. This is especially true when interpreting the findings that exclude "I don't know" responses (Table A5). Given this, some of the effects of age group observed may not be reliable. We are still actively recruiting and testing and aim to have a final sample size of 210 participants (*n*=42 per age group based on the preregistered power analysis). Furthermore, the scope of this master's thesis was limited to temporal accuracy for relatively recent events. Given this, our next steps are to continue to explore other factors within our data. First, we will explore the effect of delay by including distant events (i.e., events that took place 2-4-years ago). This will provide better insight into the effects of delay on temporal memory accuracy. Further, we will include an adult sample (18-22-years old) as a comparison group. This will allow us to better understand the developmental trajectory of temporal memory accuracy and compare memory performance for temporal details between children and adults. This will be especially important to decipher whether there are age-related differences in accuracy for time details or whether some time details are difficult to provide accurate responses for across age. We also plan to analyze the justifications for children's responses to the temporal questions. This will be important to better understand how children across different ages may go about reconstructing temporal information (Friedman, 1993). We will score justifications using Pathman, Larkina and colleagues' (2013) protocol, which will allow us to examine what types of justifications are being provided to rationalize temporal estimates and whether they differ based on the type of temporal question, how old the child is and the valence of the event. We will also analyze confidence ratings provided by participants and examine how confidence may be affected by accuracy of temporal estimates. Finally, we will further analyze the valence of the nominated and recalled events. Although parents and caregivers were asked to nominate positive and negative events, we did not provide any other specifications about *how* positive or negative the events must be. Literature on valenced events report different memory performances based on the emotional intensity of the event (e.g., Berliner et al., 2003). Given this, we plan to further differentiate the valence of events by placing events into more specific valence categories (e.g., very negative to very positive, or on a continuum of valence values). This will allow us to investigate how intensity of a valenced event may affect subsequent memory for the event.

In summary, the current study will be the first to use past autobiographical events to examine temporal memory accuracy for a multitude of temporal details across a wide range of age groups, including adults, and examining the effects of delay and valence. The findings for this study will not only contribute to the literature on children's memory for past events in a novel way, but it will also provide insight for legal practitioners on ways to best interpret legal testimonies provided by children as young as 4-years-old. Seeing that children are required to provide legal testimonies, findings from this study will aid in efforts to better understand these testimonies and contribute to the advocacy for children within the judicial system.

References

Andrews, S. J., Ahern, E. C., & Lamb, M. E. (2017). Children's uncertain responses when testifying about alleged sexual abuse in Scottish courts. *Behavioral Sciences & the Law, 35*(3), 204–224. https://doi.org/10.1002/bsl.2286

Bala, N., Lee, J., & McNamara, E. (2001). Children as witnesses: Understanding their capacities, needs, and experiences. *Journal of Social Distress and the Homeless*, *10*(1), 41–68. https://doi.org/10.1023/A:1009429602266

Bar-Haim, Y., Kerem, A., Lamy, D., & Zakay, D. (2009). When time slows down: The influence of threat on time perception in anxiety. *Cognition and Emotion*, *24*(2), 255–263. https://doi.org/10.1080/02699930903387603

Bauer, P. J. (2014). *Remembering the times of our lives: Memory in infancy and beyond*. Taylor and Francis.

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. https://doi.org/10.1111/j.1467-9280.2007.01999.x

Bauer, P. J., Wenner, J. A., Dropik, P. L., & Wewerha, S. S. (2020). *Parameters of remembering and forgetting in the transition from infancy to early childhood*.

Ben-Shachar, M., Lüdecke, D., Makowski, D. (2020). effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software,* **5**(56), 2815. doi: 10.21105/joss.02815

Berliner, L., Hyman, I., Thomas, A., & Fitzgerald, M. (2003). Children's memory for trauma and positive experiences. *Journal of Traumatic Stress*, *16*(3), 229–236. https://doi.org/10.1023/A:1023787805970

Block, R. A., Zakay, D., & Hancock, P. A. (1999). Developmental changes in human duration judgments: A meta-analytic review. *Developmental Review, 19*(1), 183–211. https://doi.org/10.1006/drev.1998.0475

Brainerd, C. J., & Reyna, V. F. (2012). Reliability of children's testimony in the era of developmental reversals. *Developmental Review: DR*, *32*(3), 224–267. https://doi.org/10.1016/j.dr.2012.06.008

Brewer, N., & Douglass, A. B. (Eds.). (2019). *Psychological science and the law*. The Guilford Press.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, *5*(1), 73–99. https://doi.org/10.1016/0010-0277(77)90018-X

Bruck, M., Hembrooke, H., & Ceci, S. J. (1997). Children's reports of pleasant and unpleasant events.

In D. Read & S. Lindsay (Eds.), *Recollections of trauma: Scientific research and clinical practice* (pp. 199–219). New York: Plenum.

Burns, P., Russell, J., & Russell, C. (2016). Children's binding of what-where-when- *who* in episodic memory: Identifying oneself in the past. *Journal of Cognition and Development*, *17*(2), 300–319. https://doi.org/10.1080/15248372.2015.1048864

Cahill, L., & McGaugh, J. L. (1995). A novel demonstration of enhanced memory associated with emotional arousal. *Consciousness and Cognition*, *4*(4), 410–421. https://doi.org/10.1006/ccog.1995.1048

Canada, K. L., Pathman, T., & Riggins, T. (2020). Longitudinal development of memory for temporal order in early to middle childhood. *The Journal of Genetic Psychology*, *181*(4), 237–254. https://doi.org/10.1080/00221325.2020.1741504

Carver, L. J. (1999). When the event is more than the sum of its parts: 9-month-olds' long-term ordered recall. *Memory*, *7*(2), 147–174. https://doi.org/10.1080/741944070

Cliver, K. G., Gregory, D. F., Martinez, S. A., Mitchell, W. J., Stasiak, J. E., Reisman, S. S., Helion, C., & Murty, V. P. (2024). Temporal memory for threatening events encoded in a haunted house. *Cognition and Emotion*, 1–17. https://doi.org/10.1080/02699931.2024.2338962

Comblain, C., D'Argembeau, A., & Van Der Linden, M. (2005). Phenomenal characteristics of autobiographical memories for emotional and neutral events in older and younger adults. *Experimental Aging Research*, *31*(2), 173–189. https://doi.org/10.1080/03610730590915010

D'Argembeau, A., Comblain, C., & Van Der Linden, M. (2003). Phenomenal characteristics of autobiographical memories for positive, negative, and neutral events. *Applied Cognitive Psychology*, *17*(3), 281–294. https://doi.org/10.1002/acp.856

D'Argembeau, A., & Van Der Linden, M. (2005). Influence of emotion on memory for temporal information. *Emotion*, *5*(4), 503–507. https://doi.org/10.1037/1528-3542.5.4.503

Deker, L., & Pathman, T. (2021). Did I visit the polar bear before the giraffe? Examining memory for temporal order and the temporal distance effect in early to middle childhood. *Applied Cognitive Psychology*, *35*(3), 785–794. https://doi.org/10.1002/acp.3804

Droit-Volet, S., Brunot, S., & Niedenthal, P. M. (2004). Perception of the duration of emotional events. *Cognition and Emotion, 18*(6), 849–858. https://doi.org/10.1080/02699930341000194

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in cognitive sciences*, *11*(12), 504–513. https://doi.org/10.1016/j.tics.2007.09.008

Earhart, B., La Rooy, D. J., Brubacher, S. P., & Lamb, M. E. (2014). An examination of "don't know" responses in forensic interviews with children. *Behavioral Sciences & the Law, 32*(6), 746–761. https://doi.org/10.1002/bsl.2141

Edgin, J. O., Spanò, G., Kawa, K., & Nadel, L. (2014). Remembering things without context: development matters. *Child development*, *85*(4), 1491–1502. https://doi.org/10.1111/cdev.12232

Fivush, R., Gray, J. T., & Fromhoff, F. A. (1987). Two-year-olds talk about the past. *Cognitive Development, 2*(4), 393–409. https://doi.org/10.1016/S0885-2014(87)80015-1

Fivush, R., Hazzard, A., McDermott Sales, J., Sarfati, D., & Brown, T. (2003). Creating coherence out of chaos? Children's narratives of emotionally positive and negative events. *Applied Cognitive Psychology*, *17*(1), 1–19. https://doi.org/10.1002/acp.854

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

Friedman, W. J. (2004). Time in autobiographical memory. *Social Cognition*, *22*(5), 591–605. https://doi.org/10.1521/soco.22.5.591.50766

Friedman, W. J., & Laycock, F. (1989). Children's analog and digital clock knowledge. *Child Development, 60*(2), 357–371. https://doi.org/10.2307/1130982

Friedman, W. J., & Lyon, T. D. (2005). Development of temporal-reconstructive abilities. *Child Development*, *76*(6), 1202–1216. https://doi.org/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. https://doi.org/10.1002/acp.1656

Gil, S., & Droit-Volet, S. (2012). Emotional time distortions: The fundamental role of arousal. *Cognition and Emotion*, *26*(5), 847–862. https://doi.org/10.1080/02699931.2011.625401

Goodman, G. S., Golding, J. M., & Haith, M. M. (1984). Jurors' reactions to child witnesses. *Journal of Social Issues*, *40*(2), 139–156. https://doi.org/10.1111/j.1540-4560.1984.tb01098.x

Goodman, G. S., Quas, J. A., Batterman-Faunce, J. M., Riddlesberger, M. M., & Kuhn, J. (1994). Predictors of accurate and inaccurate memories of traumatic events experienced in childhood. *Consciousness and Cognition*, *3*(3–4), 269–294. https://doi.org/10.1006/ccog.1994.1016

Goodman, G. S., Quas, J. A., Bulkley, J., & Shapiro, C. (1999). Innovations for child witnesses: A national survey. *Psychology, Public Policy, and Law*, *5*(2), 255–281. https://doi.org/10.1037/1076-8971.5.2.255

Goodman, G. S., & Schaaf, J. M. (1997). *Over a decade of research on children's eyewitness testimony: What have we learned? Where do we go from here?*

Gordon, H. M., Lyon, T. D., & Lee, K. (2014). Social and cognitive factors associated with children's secret-keeping for a parent. *Child Development*, *85*(6), 2374–2388. https://doi.org/10.1111/cdev.12301

Habermas, T., & Bluck, S. (2000). Getting a life: The emergence of the life story in adolescence. *Psychological Bulletin*, *126*(5), 748–769. https://doi.org/10.1037/0033-2909.126.5.748

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

Hayes, S. M., Nadel, L., & Ryan, L. (2007). The effect of scene context on episodic object recognition: parahippocampal cortex mediates memory encoding and retrieval success. *Hippocampus*, *17*(9), 873–889. https://doi.org/10.1002/hipo.20319

Hembacher, E., & Ghetti, S. (2014). Don't look at my answer: Subjective uncertainty underlies preschoolers' exclusion of their least accurate memories. *Psychological Science, 25*(9), 1768–1776. https://doi.org/10.1177/0956797614542273

Hirst, W., & Phelps, E. A. (2016). Flashbulb memories. *Current Directions in Psychological Science*, *25*(1), 36–41. https://doi.org/10.1177/0963721415622487

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Howell, D.C. (2011). Chi-square test: Analysis of contingency tables. International Encyclopedia of Statistical Science.

Hudson, J. A. (1993). Reminiscing with mothers and others: Autobiographical memory in young two-year-olds. *Journal of Narrative & Life History, 3*(1), 1–32. https://doi.org/10.1075/jnlh.3.1.01rem

Huttenlocher, J., Hedges, L., & Prohaska, V. (1988). Hierarchical organization in ordered domains: Estimating the dates of events. *Psychological Review*, *95*(4), 471–484. https://doi.org/10.1037/0033-295X.95.4.471

Jack, F., Friedman, W., Reese, E., & Zajac, R. (2016a). Age-related differences in memory for time, temporal reconstruction, and the availability and use of temporal landmarks. *Cognitive Development*, *37*, 53–66. https://doi.org/10.1016/j.cogdev.2015.12.003

Jack, F., Friedman, W., Reese, E., & Zajac, R. (2016b). Age-related differences in memory for time, temporal reconstruction, and the availability and use of temporal landmarks. *Cognitive Development*, *37*, 53–66. https://doi.org/10.1016/j.cogdev.2015.12.003

Jeunehomme, O., & D'Argembeau, A. (2018). The time to remember: Temporal compression and duration judgements in memory for real-life events. Quarterly Journal of Experimental Psychology, 72(4), 930-942. https://doi.org/10.1177/1747021818773082

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology. General*, *117*(4), 371–376.

Jones, D. P. H., Psych., M. R. C., & Krugman, R. D. (1986). Can a three-year-old child bear witness to her sexual assault and attempted murder? *Child Abuse & Neglect*, *10*(2), 253–258. https://doi.org/10.1016/0145-2134(86)90086-4

Kensinger, E. A. (2009). Remembering the details: effects of emotion. *Emotion Review*, *1*(2), 99–113. https://doi.org/10.1177/1754073908100432

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180. https://doi.org/10.3758/BF03195800

Koss, M. P., Figueredo, A. J., Bell, I., Tharan, M., & Tromp, S. (1996). Traumatic memory characteristics: a cross-validated mediational model of response to rape among employed women. *Journal of abnormal psychology*, *105*(3), 421–432. https://doi.org/10.1037//0021-843x.105.3.421

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B. (2017). “lmerTest Package: Tests in Linear Mixed Effects Models.” *Journal of Statistical Software*, **82**(13), 1-26. https://doi.org/10.18637/jss.v082.i13

Lamb, M. E., Brown, D. A., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2018). *Tell me what happened: Questioning children about abuse* (2nd ed.). Wiley-Blackwell.

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*(11–12), 1201–1231. https://doi.org/10.1016/j.chiabu.2007.03.021

Lampinen, J. M., & Smith, V. L. (n.d.). *The incredible (and sometimes incredulous) child witness: Child eyewitnesses' sensitivity to source credibility cues*.

Lee, J. K., Wendelken, C., Bunge, S. A., & Ghetti, S. (2016). A time and place for everything: developmental differences in the building blocks of episodic memory. *Child development*, *87*(1), 194–210. https://doi.org/10.1111/cdev.12447

Lenth R (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.10.6, https://CRAN.R-project.org/package=emmeans

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. https://doi.org/10.1111/j.1540-4560.2006.00489.x

MacDonald, S., Uesiliana, K., & Hayne, H. (2000). Cross-cultural and gender differences in childhood amnesia. *Memory*, *8*(6), 365–376. https://doi.org/10.1080/09658210050156822

Malloy, L. C., Lyon, T. D., & Quas, J. A. (2007). Filial dependency and recantation of child sexual abuse allegations. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*(2), 162–170. https://doi.org/10.1097/01.chi.0000246067.77953.f7

Malloy, L. C., & Mugno, A. P. (2016). Children's recantation of adult wrongdoing: An experimental investigation. *Journal of Experimental Child Psychology*, *145*, 11–21. https://doi.org/10.1016/j.jecp.2015.12.003

Malloy, L. C., Mugno, A. P., Rivard, J. R., Lyon, T. D., & Quas, J. A. (2016). Familial Influences on Recantation in Substantiated Child Sexual Abuse Cases. *Child Maltreatment*, *21*(3), 256–261. https://doi.org/10.1177/1077559516650936

McGough, L. S. (1994). Child witnesses: Fragile voices in the American legal system. Yale University Press.

McWilliams, K., Lyon, T. D., & Quas, J. A. (2019). Maltreated children's ability to make temporal judgments using a recurring landmark event. *Journal of Interpersonal Violence*, *34*(4), 873–883. https://doi.org/10.1177/0886260516645812

McWilliams, K., Williams, S., Stolzenberg, S. N., Evans, A. D., & Lyon, T. D. (2021). Don't know responding in young maltreated children: The effects of wh- questions type and enhanced interview instructions. *Law and Human Behavior, 45*(2), 124–137. https://doi.org/10.1037/lhb0000404

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about *Challenger*. In E. Winograd & U. Neisser (Eds.), *Affect and Accuracy in Recall* (1st ed., pp. 9–31). Cambridge University Press. https://doi.org/10.1017/CBO9780511664069.003

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

Orbach, Y., & Lamb, M. E. (2007). Young children's references to temporal attributes of allegedly experienced events in the course of forensic interviews. *Child Development*, *78*(4), 1100–1120. https://doi.org/10.1111/j.1467-8624.2007.01055.x

Ornstein, P. A. (1995). Children's long-term retention of salient personal experiences. *Journal of Traumatic Stress*, *8*(4), 581–605. https://doi.org/10.1002/jts.2490080405

Pathman, T., Deker, L., Coughlin, C., & Ghetti, S. (2022). Examining temporal memory and flexible retrieval of conventional time knowledge across middle to late childhood. *Journal of Cognition and Development*, *23*(4), 571–589. https://doi.org/10.1080/15248372.2022.2072846

Pathman, T., & Ghetti, S. (2014). The eyes know time: A novel paradigm to reveal the development of temporal memory. *Child Development*, *85*(2), 792–807. https://doi.org/10.1111/cdev.12152

Pathman, T., Larkina, M., Burch, M. M., & Bauer, P. J. (2013). Young children's memory for the times of personal past events. *Journal of Cognition and Development*, *14*(1), 120–140. https://doi.org/10.1080/15248372.2011.641185

Peterson, C. (2012). Children's autobiographical memories across the years: Forensic implications of childhood amnesia and eyewitness memory for stressful events. *Developmental Review*, *32*(3), 287–306. https://doi.org/10.1016/j.dr.2012.06.002

Peterson, C., & Bell, M. (1996). Children's memory for traumatic injury. *Child Development*, *67*(6), 3045–3070.

Peterson, C., Wang, Q., Hallett, D., Bartlett, S., Ma, A. Y. F., MacKay, M. M., Pretty, E. J., Viscarra, L., & Wang, E. Y. (2024). Dating early memories: When did events really happen? *Cognitive Development*, *71*, 101489. https://doi.org/10.1016/j.cogdev.2024.101489

Picard, L., Cousin, S., Guillery-Girard, B., Eustache, F., & Piolino, P. (2012). How do the different components of episodic memory develop? Role of executive functions and short-term feature-binding abilities. *Child Development*, *83*(3), 1037–1050. https://doi.org/10.1111/j.1467-8624.2012.01736.x

Porter, S., & Birt, A. R. (2001). Is traumatic memory special? A comparison of traumatic memory characteristics with memory for other emotional life experiences. *Applied Cognitive Psychology*, *15*(7). https://doi.org/10.1002/acp.766

Pouthas, V. (1993). Ontogenesis of temporal learning in the child experimental evidence and perspectives. *Psychologica Belgica, 33*(2), 171–183.

Powell, M., Roberts, K., & Guadagno, B. (2007). Particularisation of child abuse offences: Common problems when questioning child witnesses. *Current Issues in Criminal Justice*, *19*(1), 64–74. https://doi.org/10.1080/10345329.2007.12036415

Quas, J. A., Goodman, G. S., Bidrose, S., Pipe, M.-E., Craw, S., & Ablin, D. S. (1999). Emotion and memory: Children's long-term remembering, forgetting, and suggestibility. *Journal of Experimental Child Psychology*, *72*(4), 235–270. https://doi.org/10.1006/jecp.1999.2491

R Core Team. (2024). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

R. v. B.(G.), [1990] 2 S.C.R. 30 (Can.).

Richardson, P.J. (2006). Criminal pleading, evidence and practice. London: Sweet and Maxwell.

Rodriguez Steen, L. A., & Malloy, L. C. (2023). Encouraging children's clarification requests with "i don't understand" rule reminders. *Applied Cognitive Psychology.* Advance online publication. https://doi.org/10.1002/acp.4157

Scales, M. L., & Pathman, T. (2021). Flexible retrieval of semantic knowledge predicts temporal memory, but not memory for other types of context, in 4-6-year-olds. *Cognitive Development*, *59*, 101080. https://doi.org/10.1016/j.cogdev.2021.101080

Scarf, D., Boden, H., Labuschagne, L. G., Gross, J., & Hayne, H. (2017). "What" and "where" was when? Memory for the temporal order of episodic events in children. *Developmental Psychobiology*, *59*(8), 1039–1045. https://doi.org/10.1002/dev.21553

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2016). *afex: Analysis of Factorial Experiments*. R package version 0.16-1. https://CRAN.R-project.org/package=afex

State v. Taylor, 116 P. 3d 360 (Utah 2005).

St Jacques, P. L., & Levine, B. (2007). Ageing and autobiographical memory for emotional and neutral events. *Memory (Hove, England)*, *15*(2), 129–144. https://doi.org/10.1080/09658210601119762

Suengas, A. G., & Johnson, M. K. (1988). Qualitative effects of rehearsal on memories for perceived and imagined complex events. *Journal of Experimental Psychology: General*, *117*(4), 377–389. https://doi.org/10.1037/0096-3445.117.4.377

Van Abbema, D., & Bauer, P. (2005). Autobiographical memory in middle childhood: Recollections of the recent and distant past. *Memory*, *13*(8), 829–845. https://doi.org/10.1080/09658210444000430

Walker, W. R., Skowronski, J. J., Gibbons, J. A., Vogl, R. J., & Ritchie, T. D. (2009). Why people rehearse their memories: frequency of use and relations to the intensity of emotions associated with autobiographical memories. *Memory (Hove, England)*, *17*(7), 760–773. https://doi.org/10.1080/09658210903107846

Wandrey, L., Lyon, T. D., Quas, J. A., & Friedman, W. J. (2012). Maltreated children's ability to estimate temporal location and numerosity of placement changes and court visits. *Psychology, Public Policy, and Law*, *18*(1), 79–104. https://doi.org/10.1037/a0024812

Wang, Q., & Peterson, C. (2014). Your earliest memory may be earlier than you think: Prospective studies of children's dating of earliest childhood memories. *Developmental Psychology*, *50*(6), 1680–1686. https://doi.org/10.1037/a0036001

Wang, Q., & Peterson, C. (2016). The fate of childhood memories: children postdated their earliest memories as they grew older. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.02038

Wardell, V., Madan, C. R., Jameson, T. J., Cocquyt, C. M., Checknita, K., Liu, H., & Palombo, D. J. (2021). How emotion influences the details recalled in autobiographical memory. *Applied Cognitive Psychology*, *35*(6), 1454–1465. https://doi.org/10.1002/acp.3877

Waterman, A. H., Blades, M., & Spencer, C. P. (2000). Do children try to answer nonsensical questions? *British Journal of Developmental Psychology, 18*(2), 211–225. https://doi.org/10.1348/026151000165652

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag New York. https://ggplot2.tidyverse.org

Wixted, J. T. (2004). On Common Ground: Jost's (1897) law of forgetting and Ribot's (1881) law of retrograde amnesia. *Psychological Review*, *111*(4), 864–879. https://doi.org/10.1037/0033-295X.111.4.864

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science*, *2*(6), 409–415. https://doi.org/10.1111/j.1467-9280.1991.tb00175.x

Yanagida T (2024). *misty: Miscellaneous Functions 'T. Yanagida'*. R package version 0.6.6. https://CRAN.R-project.org/package=misty

Yarmey, A. D., & Jones, H. T. (1982). Police awareness of the fallibility of eyewitness identification. *Canadian Police College Journal, 6(*3), 113-124.

Yarmey, A. D., & Jones, H. P. T. (1983). Accuracy of memory of male and female eyewitnesses to a criminal assault and rape. *Bulletin of the Psychonomic Society*, *21*(2), 89–92. https://doi.org/10.3758/BF03329962

Appendix A: Supplementary Tables

Table A1

Temporal Questions and Scoring Protocol

Temporal question	Answer options or Units of measurement	Chance performance boundary	Dichotomous accuracy	Range of deviation scores
How old were you during this event?	In years (e.g., 8 years old)	Not applicable	0 = incorrect; 1 = correct	In months (e.g., 8 years old - 7 years old = 12 months difference)
What grade were you in during this event?	In grade (e.g., Grade 3)	Not applicable	0 = incorrect; 1 = correct	In grades (e.g., Grade 3 - Grade 2 = 1 grade difference)
What season was it during this event?	Winter, Spring, Summer, Fall	25%	0 = incorrect; 1 = correct	0-2
What month was it during this event?	January, February, March, April, May, June, July, August, September, October, November, December	8.33%	0 = incorrect; 1 = correct	0-6
What day of the week was it during this event?	Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday	14.29%	0 = incorrect; 1 = correct	0-3
What time of day was it during this event?	Morning, Afternoon, Evening, Night	25%	0 = incorrect; 1 = correct	0-2
What time was it on the clock during this event?	Hour and/or minute, AM or PM (e.g., 10:30AM)	Not applicable	0 = not within start and end boundary; 1 = within start and end boundary	In minutes (e.g., 10:30AM to 12:30PM interval - 9:30AM = 60 minutes difference)
From when the event started, until it ended, how long was this event?	In hours, minutes and/or seconds (e.g., 1 hour)	Not applicable	0 = incorrect; 1 = correct	In minutes (e.g., 2 hours - 1 hour = 60 minutes difference)

Note. Dichotomous accuracy was scored for exact response-match to answer. For instance, if duration estimate was 2 hours, but correct duration was 2 hours and 15 minutes, their dichotomous accuracy for duration will be scored as 0

Table A2

Outline of RStudio Functions and Packages Used for Analyses and Visualizations

Analyses	**Function**	**Package and Reference**
Mixed effects logistic regression	glmer()	lmerTest (Kuznetsova et al., 2017)
Logistic regression	glm()	stats (R Core Team, 2024)
Odds ratio (Mixed effects logistic regression)	pairs()	emmeans (Lenth, 2024)
Confidence intervals (Mixed effects logistic regression)	confint()	emmeans (Lenth, 2024)
Mixed effects ANOVA	aov_car()	afex (Singmann et al., 2024)
Effect size(η^2 or partial η^2)	eta_squared()	effectsize (Ben-Shachar et al., 2020)
Simple linear regression	lm()	stats (R Core Team, 2024)
Tukey's pairwise comparisons	pairs()	emmeans (Lenth, 2024)
Chi-squared test for goodness-of-fit	chisq.test()	stats (R Core Team, 2024)
Data visualization	ggplot()	ggplot2 (Wickham, 2016)
Descriptive statistics	descript()	misty (Yanagida, 2024)

Table A3

Delay of Events for Age Group and Valence

Type of Event (Valence)	**Age Group**	**M**	**SD**	**Mdn**	**Range**
Recent Positive	4-6-year-old	8.3	5.93	5.36	2.07, 23.39
	7-9-year-old	9.93	6.41	5.62	1.12, 21.32
	10-12-year-old	9.03	5.74	8.91	1.31, 22.64
	13-15-year-old	10.46	7.08	7.79	2.23, 23.79
Recent Negative	4-6-year-old	6.44	3.92	5.96	2.66, 18.33
	7-9-year-old	8.47	6.37	8.16	1.91, 20.6
	10-12-year-old	9.77	5.11	6.90	1.91, 22.11
	13-15-year-old	8.78	5.09	7.69	2.69, 21.19

Note. the unit of mean, median, standard deviation and range values are in months.

Table A4

Number of Children (and Percentage of Age Group) That Required Cues

Type of Event (Valence)	Age Group	No cues needed	WHAT-HAPPENED cue needed	WHO cue needed	Both cues needed
Recent Positive	4-6-year-old (*n*=19)	17 (89.5%)	2 (10.5%)	0	0
	7-9-year-old (*n*=32)	31 (97%)	0	0	1 (3%)
	10-12-year-old (*n*=44)	42 (95%)	2 (5%)	0	0
	13-15-year-old (*n*=26)	25 (96%)	1 (4%)	0	0
Recent Negative	4-6-year-old (*n*=19)	17 (85%)	1 (5%)	0	1 (5%)
	7-9-year-old (*n*=32)	28 (87.5%)	1 (3%)	1 (3%)	2 (6.25%)
	10-12-year-old (*n*=44)	39 (89%)	3 (7%)	1 (2%)	1 (2%)
	13-15-year-old (*n*=26)	23 (85%)	2 (7%)	0	1 (4%)

Table A5

Sample Sizes Excluding and Including Participants Who Said "I don't know" per Age Group and Temporal Question

Age Group	Response	Temporal Questions							
		Age estimate	*Grade estimate*	*Season estimate*	*Month estimate*	*Day estimate*	*Time of day estimate*	*Time on clock estimate*	*Duration estimate*
4-6-year-old	Sample size excluding IDK responses *(n)*	17	12	16	8	6	16	4	13
	Sample size including IDK responses *(n)*	19	19	19	19	19	19	19	19
7-9-year-old	Sample size excluding IDK responses *(n)*	31	30	29	23	21	30	19	30
	Sample size including IDK responses *(n)*	32	32	32	32	32	32	32	32
10-12-year-old	Sample size excluding IDK responses *(n)*	42	40	40	37	33	42	35	41
	Sample size including IDK responses *(n)*	44	44	44	44	44	44	44	44
13-15-year-old	Sample size excluding IDK responses *(n)*	26	25	25	24	18	26	24	24
	Sample size including IDK responses *(n)*	26	26	26	26	26	26	26	26

Note. Sample size including IDK responses match the final sample size.

Table A6

Directional Deviation Accuracy Per Age Group, Valence and Temporal Question.

Age Group	Accuracy Scale	Temporal Questions					
		Age estimate		*Grade estimate*		*Duration estimate*	
		Recent Negative	Recent Positive	Recent Negative	Recent Positive	Recent Negative	Recent Positive
4-6-year-old (n=19)	Directional deviation scores (in units of temporal question)	2.53 (9.95)	-2 (6.17)	0.14 (1.17)	-0.19 (0.4)	41.68 (348.48)	144.1 (288.04)
7-9-year-old (n=32)	Directional deviation scores (in units of temporal question)	-1.16 (8.4)	0.39 (5.78)	-0.17 (0.59)	0.03 (0.48)	-23.27 (328.9)	-66.94 (146.23)
10-12-year-old (n=44)	Directional deviation scores (in units of temporal question)	-0.29 (6.21)	0 (4.48)	-0.17 (0.5)	-0.02 (0.47)	-106.17 (307.86)	-22.02 (203.22)
13-15-year-old (n=26)	Directional deviation scores (in units of temporal question)	-0.89 (4.62)	0.46 (2.35)	0.04 (0.34)	0 (0.29)	105.7 (704.91)	-45.37 (101.99)

Note. Deviation scores are in the units of the temporal question: in months for age estimates, in grades for grade estimates, in seasons for season estimates, in days for day estimates, in time-of-day estimates, in minutes for time-on-clock estimates, in minutes for duration estimates. Standard deviations are reported in parentheses for all deviation scores.

Appendix B: Supplementary Figures

Figure B1

Age Distribution by Gender

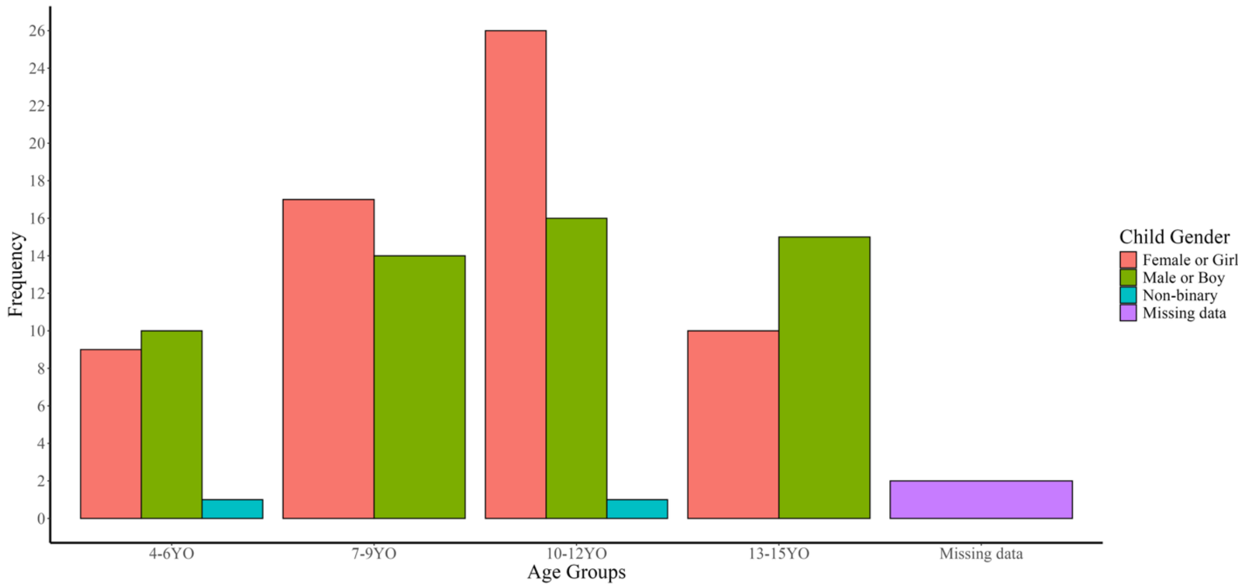


Figure B2

Number of Unusable Events and Why They Are Not Useable for Age Group Across Valence

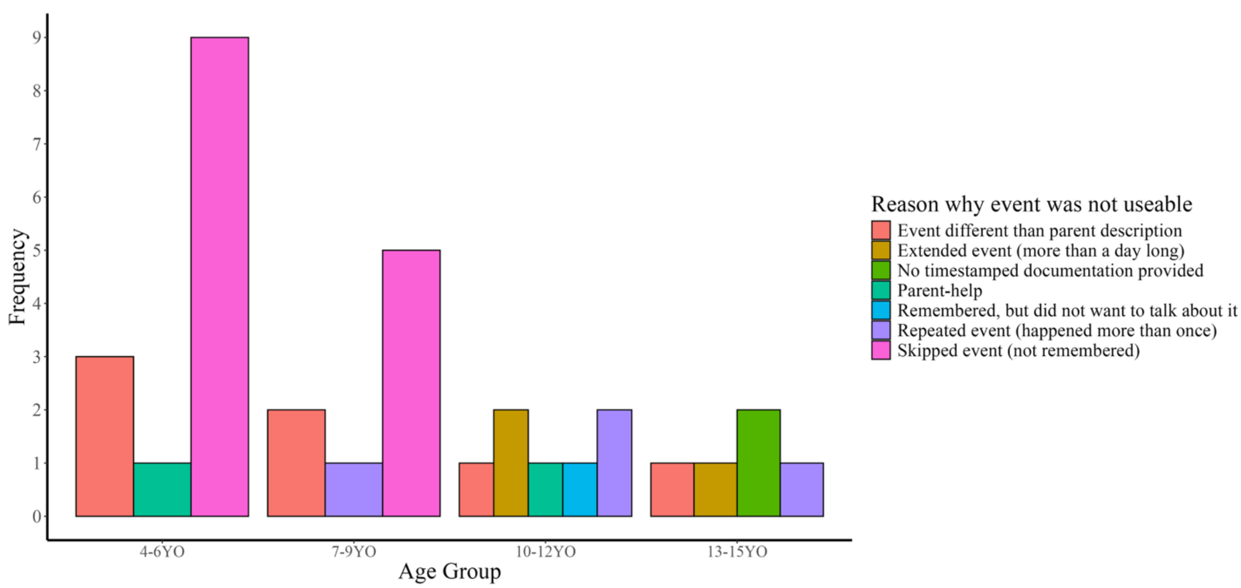


Figure B3

Overall Grade Distribution of Sample

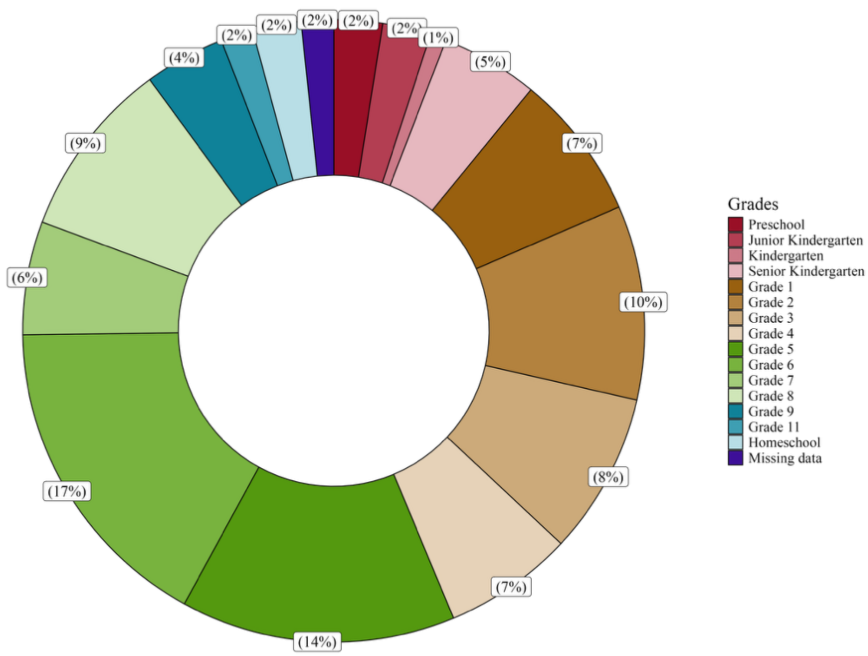


Figure B4

Overall Ethnic Distribution of Sample

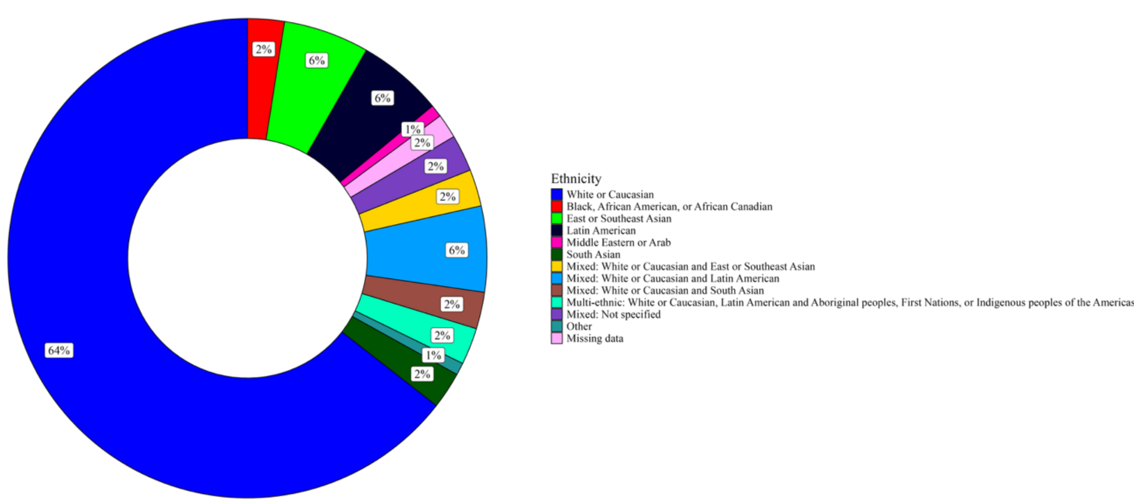


Figure B5

Overall Family Income of Sample

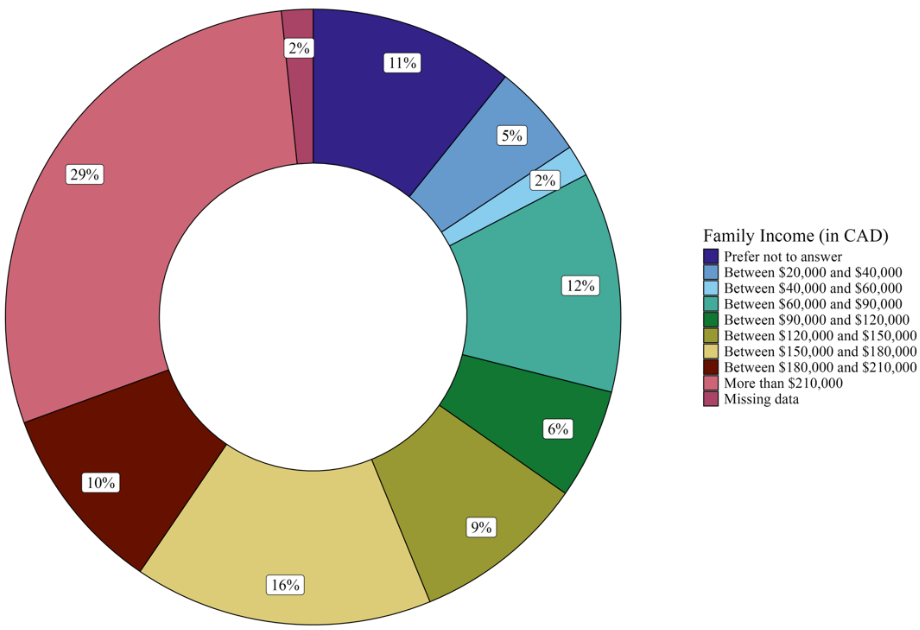


Figure B6

Overall Parental Education of Sample

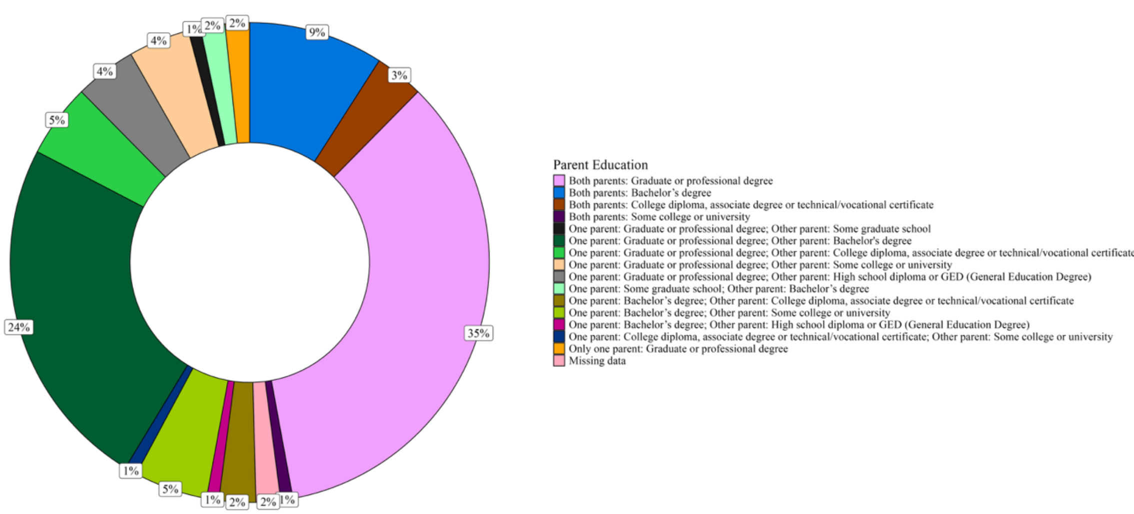
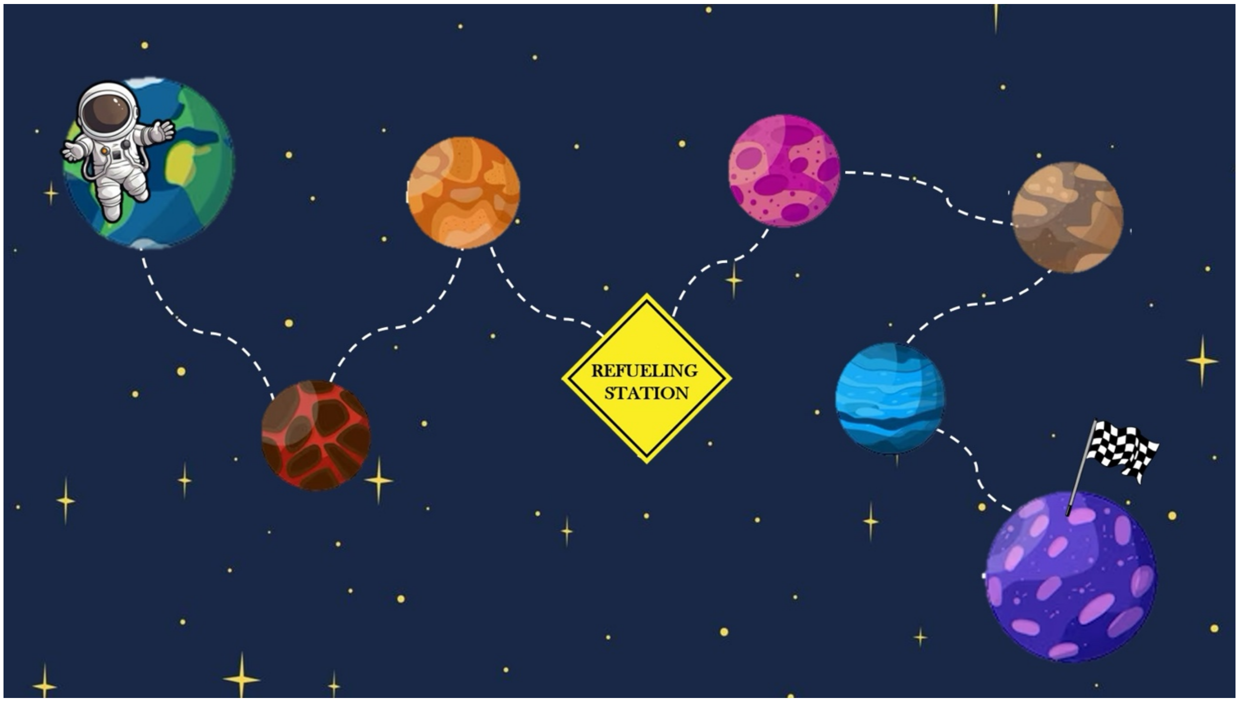


Figure B7

Progress Tracker for Main Study Session



Note. Character depicted here is the astronaut. Other options included a rocket ship and an alien.