“Lockdown was 100 months ago”: Children’s Ability to Recall Temporal Information

By

Alexandra N. Sutherland

A thesis submitted to Te Herenga Waka—Victoria University of Wellington in fulfilment of the requirements for the degree of Master of Science in Forensic Psychology

Te Herenga Waka—Victoria University of Wellington

2022

Abstract

Sometimes children need to be interviewed about a crime that they've witnessed or experienced. Investigators will ask them about what happened and who was there, and they will also want to know about *when*. However, we know little about how well children can recall temporal details of their experiences. We examined children's reporting of temporal information in two interviews with children from across New Zealand, about their experiences with COVID-19 in 2020. Results indicated that children provided more temporal information when specifically prompted than in their spontaneous accounts, but the accuracy of these details (irrespective of whether they were freely recalled or prompted) was low. Additionally, older children reported more temporal information and were more accurate than younger children. Our findings suggest that children 1) may not encode temporal details of their experiences, and 2) may attempt to offer up such information when asked, in the absence of knowledge or recall of the information. Although the judicial system may expect such information to be included in children's testimony about events—meaning that interviewers are encouraged to elicit it from children—the information is unlikely to be accurate and may undermine children's credibility or the validity of verdicts reached in the courtroom.

Acknowledgements

First and foremost, this thesis would not have been possible without my incredible supervisor Dr Deirdre Brown. Your invaluable guidance, feedback, and encouragement gave me the confidence I needed to succeed with this massive project, and I am truly grateful to have learned from the very best. Thank you so much. I'd also like to thank Dr Heather Price and Team Canada—Ezra and Adrianna—for their support and ideas throughout the process.

A huge thank you to everyone in the Applied Developmental Psychology Lab, particularly all of the wonderful volunteers and research assistants that helped me by transcribing the interviews—Danielle, Georgia P, Georgia W, Lucy, Miro, Sylvia, and Tui. An extra special thank you to Kelly and Georgia P who endured the reliability coding with me. Somehow we managed to make it out the other side (mostly) in one piece, and I really appreciate your efforts and contributions to the project.

Thank you to my MSc FPSY cohort. You were always ready to have a yarn with me, whether it was bouncing coding ideas off each other or calculating how many times we could listen to Weezer's discography on an Easterfield elevator trip to the moon (10,923). I am so grateful to have met such a kind and smart group of people. A special shout out to my office buddy Ashlee who helped me to keep VicBooks in business via our cà phê đá purchases.

Thank you to my brilliant friends who proofread the final draft and made sure it was in tip top shape—Bronwen, Catherine, Daniela, Preston, Sophia, and Sylvia. You guys rock.

Finally, thank you to my friends and family scattered across Wellington, Auckland, Napier, and Melbourne. You were never slow to celebrate my wins and always quick to support me when the going got tough. This has been an incredibly difficult journey and I am so grateful to have had such a beautiful support team behind me the entire way. The biggest thanks of all goes to Mum, Dad, Michael, Jackson, and Orla—your hugs are legendary.

Table of Contents

Abstract i

Acknowledgements ii

Table of Contents iii

List of Tables vi

List of Figures vii

List of Appendices viii

Introduction 1

Child Maltreatment in New Zealand 1

How do Jurors Evaluate a Child's Credibility? 2

The Credibility of Children 3

Beliefs about the Amount of Detail 3

What is Temporal Information, and Why is it Important? 4

Children's Development and Understanding of Time 5

Children and Temporal Information in the Courtroom 6

Age 7

Duration 7

Frequency 8

Landmarks 9

Sequencing 9

Time 10

Spontaneous and Prompted Temporal Information 11

The Effects of Delay 12

The Missing Link in the Literature 13

The Current Study 13

Hypotheses 14

Method 15

Time 1 (2020) 15

Participants 15

Procedure 16

Time 2 (2021) 19

Participants 19

Procedure 20

Data Processing 24

Transcribing 24

Coding 24

Results 27

The Place of Unverifiable Information 27

Data Preparation 27

Time 1 28

Total Amount of Information 28

Total Accuracy of Information 29

Spontaneous Information 30

Prompted Information....31

Time 2 32

Total Amount of Information 32

Total Accuracy of Information 33

Spontaneous Information....34

Prompted Information....36

Spontaneous versus Prompted....38

Changes across Time....38

Discussion 39

How Much Temporal Information do Children Provide? 40

How Accurate are Children when Providing Temporal Information? 45

Future Research....49

Conclusions 50

References....51

Appendices....64

List of Tables

Table 1 *Sample Characteristics (Time 1)* 15

Table 2 *Demographic Information – Ethnicity and Region (Time 1)* 16

Table 3 *Examples of Questions asked within each Interview Section (Time 1)* 18

Table 4 *Sample Characteristics (Time 2)* 20

Table 5 *Demographic Information – Ethnicity and Region (Time 2)* 20

Table 6 *Examples of Additional Questions asked within each Interview Section (Time 2)* 23

Table 7 *Categories when Coding for Amount* 25

Table 8 *Excerpt of the Coding Manual for Accuracy* 26

Table 9 *Amount of Information Reported in Time 1* 29

Table 10 *Accuracy of Information Reported in Time 1* 30

Table 11 *Amount of Information Reported Spontaneously in Time 1* 31

Table 12 *Amount of Information Reported when Prompted in Time 1* 32

Table 13 *Amount of Information Reported in Time 2* 33

Table 14 *Accuracy of Information Reported in Time 2* 34

Table 15 *Amount of Information Reported Spontaneously in Time 2* 35

Table 16 *Accuracy of Information Reported Spontaneously in Time 2* 36

Table 17 *Amount of Information Reported when Prompted in Time 2* 37

Table 18 *Accuracy of Information Reported when Prompted in Time 2* 38

Table 19 *Amount and Accuracy of Responses to Five Questions asked at both Time 1 and Time 2* 39

List of Figures

Figure 1 *Confidence Scale Presented to Participants in Section 5* 22

List of Appendices

Appendix A – Coding Manual (Amount) 64

Appendix B – Coding Manual (Accuracy) 70

Appendix C – Unverifiable Data and Analyses 77

Introduction

A police officer interviewed ten-year-old Rosie about alleged sexual abuse that occurred some time ago. She reported information about what happened, who was there, and where it was, but the officer needed to determine the 'when' of the case. This temporal information was important to establish a timeline of events and cross check the alibis of suspects. When the officer pressed Rosie to give the specific dates that this happened to her, she did so. However, the defendant had a strong alibi for the dates Rosie gave, which was well documented by travel tickets and CCTV footage. When this was presented to the jury in court, they doubted Rosie's accuracy about other things she had said since she was so clearly wrong about this particular detail. At the conclusion of the trial, the jury found the defendant not guilty and they were released free of any charges.

How capable are children like Rosie of providing temporal information about past events? Is it developmentally realistic for investigators, lawyers, and juries to expect children to provide it? In this study we examined whether, and how, children report temporal aspects of past events, using New Zealand's COVID-19 lockdown of 2020 as the target event.

Child Maltreatment in New Zealand

The maltreatment of children is a significant issue in New Zealand. In 2021, Oranga Tamariki—Ministry for Children received 71,400 reports of concern, 52% of which progressed to assessments or investigations involving 31,500 individual children and young people (Oranga Tamariki, 2021). There were also 3,371 reports of sexual assault and related offences against children under the age of 19 years old (New Zealand Police, n.d.). The number of reports sits significantly higher at 15,201 when including reports of acts intended to cause injury (New Zealand Police, n.d.). These figures are likely to underestimate the true prevalence when considering that many instances of child maltreatment are never reported (Gilbert et al., 2009).

Interviewers investigating allegations of child maltreatment must obtain detailed reports about what happened so that decisions can be made about prosecution as well as care and protection (Poole et al., 2011). Often in these cases the children are the only witnesses, and the sole source of evidence if there is a lack of technical or forensic evidence (Phillips et al., 2012; Salmon, 2001). Therefore, it is critical that we have a good understanding of children's capabilities to provide us with complete and accurate information. As in our example with Rosie, some of the details that children are asked to talk about may make a particularly vital contribution to the progression of the case. In the first instance, it will determine whether the investigation is likely to continue and if charges are to be laid. Should the case then reach the courtroom, their testimony will influence how credible they appear to a jury and thus have direct effects on trial outcomes, such as the verdict and sentencing.

How do Jurors Evaluate a Child's Credibility?

When there is a lack of physical evidence in a criminal trial, outcomes are heavily shaped by the quality of the witness' verbal account (Phillips et al., 2012). The quality of a piece of testimony is evaluated by the members of the jury, who each hold individual perceptions and beliefs about what a high-quality account might look and sound like (Bradfield & Wells, 2000). Common sense is essential to the role of being a juror because it must be applied to make decisions about the credibility of witnesses (Friedland, 1989). Common sense, however, is an individual experience arising from individual beliefs about the world. During a trial, jurors may overlook testimony that does not align with their personal expectations of child sexual assault cases and victims (Goodman-Delahunty et al., 2010). Jury eligible citizens in Australia exhibited extensive uncertainty when questioned about children's memory and experiences of sexual abuse, which indicates that this is not common knowledge to the average person (Cossins et al., 2009). This gives space for misconceptions and assumptions to influence decision-making in these cases.

The Credibility of Children

Jurors tend to believe that children's testimony is less credible than that of adults (Bruer & Pozzulo, 2014; Klemfuss & Ceci, 2012). Ross and colleagues (2003) found that jurors perceive child witness credibility in terms of cognitive ability and honesty, and that honesty predicted the verdict of the trial. Some studies found younger children are perceived as more credible than older children (Rogers & Davies, 2007), whilst others found the opposite (McCauley & Parker, 2001). Tabak and Klettke (2014) found that although a fifteen-year-old was seen to be more competent in their recollection, they were also perceived as less trustworthy and more likely to lie when compared to a younger child. Jurors believe that children will fabricate stories or be easily manipulated into providing false reports, which undermines their credibility (Morison & Greene, 1992; Quas et al., 2005). However, jurors also identified that the naivety of young children means they may be less likely to invent a narrative of sexual victimisation (Goodman et al., 1989; Tabak & Klettke, 2014). The wide variation in perceptions of child witness credibility further emphasises the need to understand children's capabilities when providing eyewitness testimony.

Beliefs about the Amount of Detail

Many people have inaccurate beliefs about how memory works and tend to overestimate their knowledge of it (Ost et al., 2017). One such belief held by the public is that a very detailed account is likely to be more accurate than one with less detail (Akhtar et al., 2018; Quas et al., 2005). Henry and colleagues (2011) found the amount of detail provided by children was a significant predictor of their perceived credibility. The extent to which children include detail in an account may therefore activate this assumption. However, research does not support this belief and provides evidence for the opposite—that true memories are not necessarily characterised by a richness of detail, and that false memories are not necessarily sparse in detail (Arndt, 2012; Otgaar et al., 2008; Otgaar et al., 2012).

Loftus and Bernstein (2005) summarised this well when writing "just because the memory report is detailed […] does not mean it really happened" (p. 111). Improving the understanding of children's capabilities when discussing past events is essential to educate jurors and challenge inaccurate assumptions about children's testimonial accuracy.

What is Temporal Information, and Why is it Important?

Temporal information refers to details pertaining to time, such as calendar dates and clock times. Although these may aid our understanding of how an event is situated in time, temporal information encompasses a much broader range of details. For example, calendar dates and clock times may be inferred from propositional information in the environment (Linton, 1975, 1978, as cited in Estes, 1985). Temporal information also includes how events relate to each other in time, such as within a chronological order. Words such as 'then', 'next', and 'after' convey temporal meaning of a sequence of events even though, at first, they appear to have little relation to time.

The importance of temporal information is that it creates a logical timeline of events, which contributes to the coherence of a narrative (Habermas & Bluck, 2000). Coherence can powerfully influence the assessment of an eyewitness' credibility, as credibility is reduced when testimony includes incorrect or inconsistent details, even if they are trivial (Berman & Cutler, 1996; Berman et al., 1995; Borckadt et al., 2003). This means that jurors may be susceptible to doubting a witness' credibility based on one or a small number of errors, and therefore will be less likely to convict based on this evidence (Puddifoot, 2020). Moreover, memory reports may be considered less believable if they are not corroborated by the reports of other witnesses or the physical evidence of the case (Goodman et al., 1987). This was seen in our example with Rosie when the jury decided that she could not be trusted when her dates were incongruent with the physical evidence. This can prove alarming given witnesses commonly include both accurate and inaccurate details in their reports of events. We do not

know whether the accuracy of one category of information tells us anything useful about the accuracy of other types of details. Arguments made by counsel, or beliefs held by jurors about the implications of incorrect details for the overall credibility of a witness, may significantly influence the outcome of a trial, which disproportionately impacts vulnerable witnesses and/or innocent defendants.

Children's Development and Understanding of Time

For a child witness to provide testimony in court, they must have sufficient language and speech skills so that they can comprehend court proceedings as well as present a logical and coherent account of what they remember (Carter et al., 1996; Saywitz, 2002). Children's vocabulary significantly increases as they age (Anglin et al., 1993), as does the level of detail given in a report (Lamb et al., 2000; Phillips et al., 2012). Additionally, there is evidence that their ability to form a coherent narrative improves over time (Fivush, 2011; Miragoli et al., 2017). A meta-analysis of 26 studies demonstrated that maltreated children consistently have poorer language skills than non-maltreated children (Lum et al., 2015). This suggests maltreated children may struggle to understand the questions asked of them and might find it difficult to articulate their experiences when required to testify. However not all children that get interviewed will have been maltreated, and thus it is worth noting that a child providing a clear and articulate report is not necessarily being untruthful in their testimony.

A key part of forming a coherent and chronological account is the knowledge and use of temporal information, however research suggests that children use temporal language before they understand what it means (Busby-Grant & Suddendorf, 2011; Tillman & Barner, 2015). Children use temporal language such as 'minute' and 'hour' from as young as two years old, but do not have a good understanding of their meanings until much later, at around six or seven years of age (Tillman & Barner, 2015). Busby-Grant and Suddendorf (2011) found that sequencing words 'before' and 'after' are produced more often and more

accurately by five-year-olds than three-year-olds. For example, 53% of three-year-olds used 'before' accurately and 69% used 'after' correctly, which increased to 77% and 83% accuracy respectively in five-year-olds. Similar results were seen in the use of other temporal terms such as 'yesterday', 'soon', and 'next week' (Busby-Grant & Suddendorf, 2011). This suggests an early lack of understanding of the function of the words resulting in inaccurate use, however accuracy appears to increase as the children develop. This may be attributed to the development of temporal metamemory throughout childhood—that is, the knowledge and understanding of how temporal information is remembered (Friedman, 2007). Friedman (2007) suggests that children may not be appropriately equipped to make accurate temporal judgements until at least the age of twelve, and that children without fully developed temporal metamemory may be more likely to guess. This is important to consider in a forensic setting, because if child witnesses inaccurately use temporal vocabulary their overall credibility may be compromised.

Children and Temporal Information in the Courtroom

Children are often questioned about temporal information in criminal trials so that lawyers and other court participants can determine their ability make temporal judgements, and evaluate the likelihood that their memory will be accurate overall (Friedman & Lyon, 2005). Temporal details such as the time of day or an estimation of when the allegations began convey important forensic information, and a child's ability to provide these details can make their testimony more valuable (Orbach & Lamb, 2007). Such information is also important to allow the accused to mount their defence (e.g., establish an alibi). However, Scarf and colleagues (2017) found whilst young children aged between three and six can accurately report what happened and where, they are less accurate in reporting when. This is consistent with research that suggests episodic memory for 'when' information is developed more slowly than that for 'what' and where' information (Cuevas et al., 2015; Hayne &

Imuta, 2011). With this in mind, Lyon and colleagues (2017) discuss that—as seen in our example with Rosie—lawyers will press children to provide temporal information which can undermine the credibility of children as witnesses. This suggests children may be getting pushed into providing information that they are developmentally unequipped for.

Age

A witness' testimony in court often begins by stating details about themselves, such as their name, address, and date of birth, which not only confirms the identity of the witness, but perhaps serves as the first indicator of accuracy and hence credibility. Children aged six years old are 58% and 64% accurate when giving their birthday day and month respectively, and these figures rise to 90% and 94% by the age of ten (Wandrey et al., 2012). Additionally, witnesses are routinely asked for their current age, and their age at the time of the event they are testifying about. Bauer and colleagues (2007) asked children aged seven to ten to indicate their age at the time of different events and holidays, and found they were 89% accurate in doing so. This is in contrast to research with a sample of maltreated children who were around 50% accurate in identifying their age at various foster placements and court visits (Wandrey et al., 2012). Providing accurate information about age at the time of a sexual crime is particularly important as criminal charges and their associated sentences differ depending on the age of the victim. For example, in New Zealand, the term of imprisonment for an act of sexual connection increases as the age of the victim lowers from under sixteen to under twelve years of age (Crimes Amendment Act 2005, ss. 132-134).

Duration

Witnesses may be asked to estimate how long an event occurred for. Much of the existing research on children's ability to do this has been conducted in a laboratory setting, involving comparing the duration of brief auditory tones (e.g., seconds) after short delays

(e.g., minutes; Droit-Volet, 2003; Lustig & Meck, 2011; McCormack et al., 1999). Older children at eight years old can differentiate varying durations of stimuli with more accuracy than younger children (Droit-Volet, 2003) which may be attributed to enhanced sensitivity to duration throughout development as discussed by McCormack (2015). Findings from these types of studies are not largely applicable to a forensic context – firstly due to the unnatural environment that laboratory studies occur in, but also because witnesses are not typically interviewed immediately following a crime and they won't testify in court for an even longer period (Goodman et al., 1992). This was addressed in a pivotal study by Friedman and colleagues (2010), which asked children aged six to twelve to judge the amount of time they spent at an appointment at the paediatrician's office. It was found that the children were reasonably accurate when doing this task, however the authors suggested that they would not be capable of giving the level of precision that the legal system would require.

Frequency

Another type of temporal information requested from child witnesses is how often an event happened. This is to establish the severity of the offending in terms of its frequency, which in turn determines the number of charges to be laid. Findings from Roberts and colleagues (2015) show less than 25% of children aged four to six with repeated experience of an event provided the correct number of times it occurred. Similarly, 9.4% of four- to eight-year-olds who experienced an event numerous times correctly identified how many times it occurred, although this rises to 67% when including children who were one time shy (Sharman et al, 2011). Maltreated children were very poor in recalling the exact number of foster placements and court visits experienced, with 24% and 14% doing so correctly respectively (Wandrey et al., 2012). A child's ability to recall the frequency of an event they have experienced is important as it may also affect their perceived credibility. For example, if a child states an outlandish number then jurors may dismiss their testimony as unreliable.

Landmarks

A common strategy to facilitate the retrieval of temporal information is to anchor thoughts in time by using a major memorable event, such as holidays, birthdays, and seasons. Questions may be posed about the target event's temporal position relative to that of the landmark event, and for example whether it happened before, after, or near it (McWilliams et al., 2019). Strube and Weber (1988, as cited in Friedman & Lyon, 2005) found very young children performed better at sequencing events when a landmark such as Christmas Day was provided, and Friedman (1992) saw that children as young as four years old had some capacity at recalling memories from the past and accurately linking them to landmark events including Easter and Christmas. However, a more recent study found children ranging from four to thirteen years old struggled to remember the proximity or order of a target event relative to a major holiday (Friedman & Lyon, 2005). Interestingly, a sample of maltreated six- to ten-year-old children exhibited a prospective bias in their answers when asked whether the current day, their most recent court visit, and their most recent change in foster placement was near, before, or after their birthday (McWilliams et al., 2019). This means that they tended to reference a forthcoming birthday rather than one in the past, which has implications in terms of their accuracy when answering these type of questions. Legal practitioners may think that referring a child to a landmark event will allow them to make an accurate judgement of when a target event happened, but such assumptions are risky and ill advised (Friedman & Lyon, 2005). Questions with landmarks are inherently subjective, and as such, children's responses to them may be misleading or inaccurate.

Sequencing

The ability to correctly order events as they happened is critical in establishing a timeline for a criminal case. Children can perform well when ordering events that are

unrelated to each other, with six-year-olds doing so accurately after a three month delay (Friedman & Lyon, 2005). Children's abilities to sequence events is generally better after a short, rather than a long, delay (Price & Evans, 2021). However, children interviewed in forensic contexts are likely to have experienced related events and will be expected to recall them accurately after a longer delay. By the age of five, children have a developed understanding of causality and are better at recalling the sequence of events with logical connections than those without (Fivush et al., 1992). When five- to ten-year-olds completed numerous tasks where they had to sequence everyday events such as the months on a calendar or a child's typical daily routine, it was found that the ability to do so correctly improves across development (Moore et al., 2014). This may be due to the association between the accurate ordering of past personal events and higher levels of conventional time knowledge acquired over time (Friedman et al., 2011). To better understand children's capabilities in the real world rather than in the laboratory, Pathman and colleagues (2013) had parents and children take photographs each day for four weeks and found both children and parents performed poorly when precision was required in their sequencing of the photographs. Acute precision is required when testifying in court, and therefore it is imperative that children's abilities are not overestimated.

Time

Perhaps the most instantly recognisable type of temporal information asked of child witnesses is that of specific dates and times, but the extant literature is mixed regarding their abilities to provide this correctly. There is evidence that preschool aged children can make accurate judgements about events that happened one day, one year, and three years ago (Busby-Grant & Suddendorf, 2011), and four- and five-year-olds are able to report events that happened at a specified time point (Fivush et al., 1984; Friedman, 1991). By the age of eight, children can recall some temporal details such as the day of the week and the month

better than chance (Friedman et al., 2011; Pathman et al., 2013). Around half of seven- to ten-year-olds providing narratives of personal events included at least some indication of time, and they were 92% accurate in identifying the year and season (Bauer et al., 2007). Contrastingly, memory for month and season can be limited in children as old as twelve (Friedman & Lyon, 2005), and only 10% of maltreated children aged between six and ten were able to correctly identify the month in which past events occurred (Wandrey et al., 2012). Further, Friedman and Lyon (2005) found children's performance differed depending on the type of temporal information requested of them, which may partially explain the variability in results across studies.

Spontaneous and Prompted Temporal Information

Interviewers have many different techniques that they may employ to elicit a complete and accurate account from a child. In accordance with best practice protocol, interviewers should pose open-ended questions to allow for free recall of an event before asking closed questions (Lamb et al., 2007). This is to avoid the suggestion of any content to the child. Research suggests the use of open-ended questions elicits highly accurate memory accounts regardless of age (Lamb et al., 2018), and that this method is more accurate than those that use focused prompts (Lamb et al., 2007; Lamb & Fauchier, 2001). Phillips and colleagues (2012) found that more information relevant to the investigation was elicited from children using open questions than from closed, leading, and forced choice questions. This may not always be the case, however, as focused prompts may be needed to gain important information that the child was unable to provide on their own. For example, children rarely offer temporal information spontaneously (Friedman & Lyon, 2005; Orbach & Lamb, 2007; Price & Evans, 2021). Young children in particular may struggle to provide temporal information spontaneously, but the quantity increases with age (Orbach & Lamb, 2007). Focused prompts run the risk of eliciting inaccurate information (Brown & Lamb, 2015), but

the rarity of spontaneous temporal details from children means that focused prompts may inevitably be required if investigators are seeking specific information.

The Effects of Delay

As mentioned previously, children are rarely interviewed immediately after the event occurred (Goodman et al., 1992). This is unfortunate as it is well documented that children interviewed soon after an event can recall it with more detail and with greater accuracy than children interviewed after a delay (Brown et al., 2015; Tizzard-Drover & Peterson, 2004). Delay is an important factor to consider when investigating memory because of its effects on memory retrieval processes. Long delays between event and interview are typically associated with less detail and lower accuracy in children, particularly in response to free recall questions (Hudson & Fivush, 1991; La Rooy et al., 2005; Pipe et al., 1999; Salmon & Pipe, 1997). However, others have found the opposite—that children's reports do not lose detail and may even get better over time (Bruck et al., 2002; Pipe et al., 2004). A limitation of these studies is that they were conducted in a staged environment, which means that the findings may not be easily applied to an everyday, or traumatic, context such as one that a child might be interviewed about. Peterson (2011) observed that when children reported their experience of a serious physical injury after a one- and two-year delay, they reported more details, but accuracy decreased as more time passed. This is an important finding considering the forensic context where children may be interviewed multiple times about the same event, during which new and previously unreported details emerge (La Rooy et al., 2005). These details are often less accurate than those reported consistently across interviews (Salmon & Pipe, 1997). This, along with the typically long delay since the event, may have implications for the perceived credibility of child witnesses and associated trial outcomes.

The Missing Link in the Literature

As discussed, there is little and mixed evidence regarding children's capabilities to provide temporal information. A particular gap in the literature is the lack of research done with naturally occurring events in children's lives. Studies conducted inside a laboratory setting lack the contextual richness of everyday life which contains more detail for children to remember. Research of how children recall temporal information from their regular life experiences may give us a deeper understanding of their abilities if they are required to do so in court. This will enable us to inform interviewers about whether and how they should ask questions on the matter, as well as inform jurors about what is realistic to expect from children of different ages.

The Current Study

The following study aimed to address these gaps by interviewing children about their memory of New Zealand's national lockdown that was implemented in 2020 in response to the COVID-19 pandemic. The lockdown presented a unique opportunity to ask children about a naturally occurring event that had become integrated into their everyday lives, with the ability to verify the accuracy of their responses against extensive official records. The children answered questions about the numerous temporal markers that were present throughout this time, such as the specific dates on which lockdown began and ended, the duration of the Alert Levels within the lockdown, and the chronological order that they occurred in. Also situated within the lockdown were significant public holidays including Easter weekend and ANZAC Day, which children may have used as landmarks to anchor their thoughts and determine answers to other temporal questions.

The current study investigated the amount of temporal information that children provide spontaneously and when prompted, and how accurate they are when reporting this kind of detail. Different types of temporal information will also be discussed in terms of that

most commonly provided and their respective accuracy results. The age of the children was considered to determine if there is a relationship between the amount and accuracy of temporal information and age. Differences between spontaneous and prompted information were also analysed. Finally, the children's responses in two interviews conducted approximately one year apart were compared to determine how the amount and accuracy of temporal information may vary across time.

Hypotheses

As described, there is a limited and mixed body of research to guide the formation of hypotheses but nonetheless we drew from broader eyewitness testimony literature, regarding delay and spontaneous versus prompted recall, to inform our expectations. Given the impact of delay on memory, we expected that at the second interview, known as Time 2, children would report less temporal information than they did at the first interview, labelled Time 1. However, because we added questions that specifically asked for temporal details, we predicted that we would see more of such information reported at Time 2 than Time 1. Accuracy of information tends to decline with delay and is lower during prompted than spontaneous recall. As such, we expected that overall accuracy, and accuracy during both spontaneous and prompted recall would be lower at Time 2 than at Time 1 due to the effects of delay and prompting, respectively. Given Friedman (2007) suggested that the age of twelve is a key stage in children's ability to make accurate temporal judgements, we expected that older children would be more accurate than younger children, and provide more information than younger children due to developmental differences such as the acquisition of language and greater understanding of time. We explored the frequency and accuracy of various types of temporal information provided, however we made no predictions about this due to the exploratory nature of the current study and the lack of prior literature on the topic.

Method

This study was granted ethical approval by the Te Herenga Waka—Victoria University of Wellington Human Ethics Committee under application #0000028654.

Time 1 (2020)

Participants

Parents of New Zealand children were invited to consent for their child to participate in this study. Consent was obtained for 119 children, and 97 participated in an interview (81.5%). Upon reviewing the recorded interviews, it was determined that a participant who completed a session did not clearly assent to their participation. This child was removed from the sample, resulting in a final sample size of 96 participants. Table 1 shows the participant's ages ranged from 6–14 years (*M* = 8.90, *SD* = 2.14), and there were 54 females and 42 males. The majority of participants identified as NZ European/Pākehā (82.29%), and resided in the Wellington region (92.71%), as seen in Table 2.

Table 1

Sample Characteristics

		Gender		
		Female	Male	Total
Age in years	6	6	7	13
	7	10	9	19
	8	11	4	15
	9	5	6	11
	10	10	5	15
	11	4	5	9
	12	4	5	9
	13	3	0	3
	14	1	1	2
Total		54	42	96

Table 2

Demographic Information – Ethnicity and Region

Demographics	Response	Total
Ethnicity	NZ European/Pākehā	79
	Māori	9
	Indian	5
	American	3
	African	2
	British Indian	2
	Latin American	2
	Middle Eastern	2
	Chinese	1
	Cook Island	1
	European	1
	European-Irish	1
	Samoan	1
Region	Wellington	89
	Auckland	2
	Manawatū-Whanganui	2
	Bay of Plenty	1
	Marlborough	1
	Waikato	1

Note. Ethnicity responses sum to greater than the sample size as parents/caregivers of participants were able to elect multiple ethnicities.

Procedure

Children from across New Zealand were recruited to participate in the study using advertisements in school newsletters, social media posts, and invitations sent to parents of children who had previously participated in studies in the Applied Developmental Lab. These recruitment approaches directed parents to a link where they could complete the 'Informed Consent' survey on Qualtrics, which consisted of an information sheet followed by an electronic statement of consent. The information sheet detailed the aims of the project, what would happen to the data provided by participants, and explained that the research was confidential. After viewing this information, parents indicated consent with a tick box and provided contact information for research assistants to arrange a suitable time for the

interview. Once this was arranged over email, parents were sent a link to an online Zoom meeting room approximately two days prior to their child's interview session.

Two research assistants conducted the interviews over a period of eleven weeks. The interviews followed the interview protocol for Time 1, questions from which can be seen in Table 3. The interviewers began each session by introducing themselves and what they would be talking about today, informed the children that they were allowed to stop talking whenever they liked, and that their identity would be anonymous to people outside of the research term. Parents and caregivers were asked to stay out of camera view and avoid helping their children with anything other than technical support as required. The interview itself began by asking children for their assent to participate, their date of birth, and their age both presently and when COVID-19 started. The interviewers then guided participants through three different sections, giving minimal encouragements and inviting elaboration when needed. The first section discussed what school days and weekends were normally like for the child before COVID-19 existed. The second section examined this same topic but in the present day, whilst also asking for the participant to note things that were the same or different as before COVID-19 happened. Section 3 also asked about school days and weekends, but specifically about the time period that COVID-19 began to enter their lives and the national lockdown that followed. This section also discussed the child's personal experiences with things in their life that changed because of COVID-19. Participants were then asked about how easy or difficult it was to remember what life was like at the three different time points. The interviewer closed the interview by asking questions about what the children liked and didn't like about COVID-19, when they thought that life might return to normal, and what normal life might look like. After completing their interview, each participant was sent a $10 voucher via post or email as a thank you token for their contribution to the research.

Table 3

Examples of Questions asked within each Interview Section

Interview section	Question examples
Introductory questions	"How old are you right now?"
	"When is your birthday?"
	"How old were you when COVID-19 started?"
1. Life before COVID-19	"Tell me all about what a regular school day was like"
	"Tell me all about what weekends were like"
2. Current day	"What are school days and weekends like now?"
	"What's different from before?"
	"What's the same as before?"
3. During COVID-19	"How long ago did COVID-19 start?"
	"When did the COVID-19 changes start for you, and what things changed?"
	"How did you know COVID-19 was going to be a big deal?"
	"What were school days and weekends like during lockdown?"
	"Tell me about a day that was different from your usual COVID-19 days"
	"How long did you have to stay in your bubble?"
Closing questions	"What is the biggest change in your life since COVID-19 started?"
	"Tell me about something you like and don't like that has happened since COVID-19"
	"When do you think things will get back to being like they were before COVID-19?"

Time 2 (2021)

Participants

Parents of children that participated in Time 1 were invited via email to participate in Time 2 (N = 96). Consent was obtained for 54 children and 49 of those participated in an interview (90.74%), a recruitment of 50.52% of Time 1 participants. Table 4 shows the children's ages ranged from 6–14 years (M = 9.68, SD = 2.28), and there were 29 females and 18 males. The majority of participants identified as NZ European/Pākehā (89.36%) and resided in the Wellington region (95.75%), as seen in Table 5.

The Auckland region experienced a greater quantity of days in COVID-19 lockdown than the remainder of New Zealand. Upon the commencement of recruitment for Time 2, it was decided that any participants residing in Auckland would be excluded from participation. As a result, one child was not invited to participate in another interview. Despite having this rule in place, missing address information meant that one participant from Auckland was interviewed at Time 2. This participant was later removed from the dataset for this reason, in addition to being an outlier of the sample in terms of their age.

One participant was removed from the dataset after their interview as researchers learned that they had completed the Parent/Caregiver Questionnaire alongside their parent. This questionnaire was intended to be completed by only the parents and caregivers of participants, and the child's participation in this activity may have prompted recall of memories of the target event. The contents of the Parent/Caregiver Questionnaire will not be analysed nor discussed in the current study, but it must be noted that it was the reason for this participant's removal from the study. This resulted in a final number of 47 participants for Time 2.

Table 4

Sample Characteristics

		Gender		
		Female	Male	Total
Age in years	6	1	1	2
	7	4	3	7
	8	8	4	12
	9	1	1	2
	10	2	2	4
	11	6	2	8
	12	4	1	5
	13	2	4	6
	14	1	0	1
Total		29	18	47

Table 5

Demographic Information – Ethnicity and Region

Demographics	Response	Total
Ethnicity	NZ European/Pākehā	42
	American	2
	Indian	2
	Māori	2
	Chinese	1
Region	Wellington	45
	Bay of Plenty	1
	Marlborough	1

Note. Ethnicity responses sum to greater than the sample size as parents/caregivers of participants were able to elect multiple ethnicities.

Procedure

Time 2 interviews were conducted approximately eleven months after Time 1 interviews (*M* = 327.96 days). The author conducted all of the interviews across a period of seven weeks.

Parents were emailed and asked to complete the Informed Consent survey, which was the same format as that for Time 1. When the survey was submitted, parents were contacted to arrange a suitable interview time and sent a Zoom link approximately two days before their

child's session. Parents were asked to stay out of camera view and avoid helping their children with anything other than technical support as required. Children were again required to provide verbal assent before the interview began. The interviewer then guided each child through the interview protocol, giving minimal encouragements and inviting elaboration when needed.

Interviews were conducted following the interview protocol for Time 2, which was adapted from that of Time 1. The key difference between Time 1 and Time 2 interviews was the addition of questions specifically assessing recall of temporal information, and these questions can be seen in Table 6. Interview questions from Time 1 (Table 3) were also asked to participants at Time 2. At the beginning of each interview the interviewer played a guessing game, using the introductory section's landmarking questions to try and work out the child's birthday. This was done to assess how children understand and describe time, specifically in the context of when their birthday is, in relation to Easter, Christmas, and the present day. The interviewer pretended to use the given responses to guess the child's birthday but would always do so incorrectly to avoid spoiling analysis of their accuracy when providing their birthday later on in the interview. This game was included to build upon the findings of McWilliams and colleagues (2019), but these particular analyses will not be conducted not discussed in the current study.

After the birthday guessing game, the interviewer asked the same questions from Sections 1, 2, and 3 that were asked in Time 1. These sections discussed what school days and weekends were like before and during COVID-19, as well as at the current day. Time 2 interviews included Section 4, which asked the child what school days and weekends were like when the COVID-19 lockdown first ended, and what might be the same or different about that time when compared to before COVID-19 started and the present day. This section gave the children a new time period to think about and reflect upon.

Section 5 was also a new addition to the interview protocol. Here, structured questions specifically asked about temporal information such as dates, estimates of duration, and the chronological order of the different Alert Levels within lockdown. It also contained two sets of landmarking questions about the child's birthday and Easter relative to lockdown. After each question children were asked to explain how they knew the information, and to indicate how confident they were about their answer on a visual scale depicting five glasses of water (see Figure 1). The interviewer held up a physical copy of this confidence scale to the camera, and children verbally indicated which glass of water best corresponded with their level of certainty, with 0 being the least and 4 being the most confident.

Some children gave extensive responses in the early stages of the interview. When this occurred, the interviewer skipped some questions to ensure that the temporal questions were covered, and to avoid delaying a timely start to other participants' interviews. 40 children completed all sections of the interview protocol, 6 completed all but one section, and 1 child completed fewer sections. All participants completed the temporal questions in the birthday guessing game and Section 5.

After completing their interview, each participant was sent a $10 voucher via post or email as a thank you token for their contribution to the research.

Figure 1

Confidence Scale Presented to Participants in Section 5

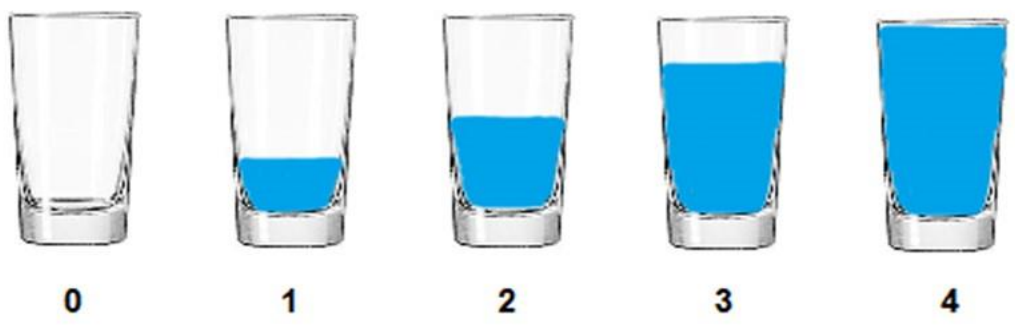

Table 6

Examples of Additional Questions asked within each Interview Section

Interview section	Question examples
Introductory questions	"Is Easter/Christmas/today near your birthday?"
	"Is Easter/Christmas/today before or after your birthday?"
	"How long ago was your last birthday?"
4. After COVID-19 lockdown	"How long ago did lockdown end?"
	"What were school days and weekends like when lockdown first ended?"
	"Tell me what is the same or different between before lockdown and when it first ended?"
	"Tell me what is the same or different between when lockdown first ended and now?"
5. Temporal questions	"How long ago was lockdown?"
	"What season was it when lockdown happened?"
	"When did lockdown start and finish, and how long did it last for?"
	"When did Alert Level 1/2/3/4 start, and how long did it last for?"
	"Was lockdown/Alert Level 1–4 as long as the summer holidays, a term of school, or the mid-year school holidays?"
	"Tell me what order that the four Alert Levels of lockdown happened in"
	"Did Easter/your birthday happen during lockdown?"
	"What Alert Level was New Zealand in when Easter/your birthday happened?"
	"Was lockdown near Easter/your birthday, and was it before or after it?"

Data Processing

Transcribing

Zoom offers an auto-transcription service for recorded video conferences, which produced automated transcripts for research assistants to use. Seven and eight research assistants worked on Time 1 and Time 2 respectively, and were required to sign confidentiality agreements before beginning work on the project. The research assistants listened to the entirety of each interview whilst checking the accuracy of the Zoom transcript and correcting any errors as they went. Factors that contributed to errors in the transcripts were how close to a microphone the speaker was, the speed of the speaker's speech, and the accent of the speaker. Typically, the transcript was more accurate for the interviewers' speech due to often wearing a small microphone and speaking slowly and clearly, whereas children generally sat far from the computer and spoke relatively quickly. Children also tended to mumble. The most common driver of errors was the New Zealand accent held by the majority of participants. The transcribing process took a mean of 159.57 minutes (*SD* = 53.29) or 2.66 hours for each of the 47 Time 2 interviews, ranging from 75 to 330 minutes. Records for Time 1 are not as robust, but data from 30 transcripts show a mean transcribing time of 134.83 minutes (*SD* = 49.38) or 2.25 hours, ranging from 70 to 265 minutes to complete.

Coding

Amount. All interviews across Time 1 and Time 2 were coded for the amount of temporal information provided by the participants. To do this, seven different categories were developed, each based on a key type of temporal information (see Table 7). These were Age, Birthday, Duration, Frequency, Landmark, Sequence, and Time. The research assistant coded a transcript by identifying pieces of temporal information and marking each of them with the corresponding code according to the definitions and rules set out in the coding manual (see Appendix A). They would also mark whether the piece of temporal information was offered

spontaneously or whether the child was prompted to produce it by the interviewer's question. For example, a date given in response to a question regarding the child's birthday would be coded as 'Prompted Birthday', whereas an estimation of how long something lasted that the child provided on their own volition would be coded 'Spontaneous Duration'. These marks were then summed for each participant to give the total overall amount of temporal information provided. Within this were the totals of both spontaneous and prompted information, and the total number of temporal details in each of the seven code types.

Table 7

Categories when Coding for Amount

Code Type	Definition	Examples of codes
Age	Number of years or months of age	Seven, 10 and a half
Birthday	Specific date of birth	23rd of February 2010
Duration	A given period of time that describes how long a defined event lasts for	Seven weeks, a long time, ten minutes
Frequency	Phrases that reference the changing or enduring frequency of an event	Every day, usually, sometimes, twice a week
Landmark	Specific points of reference that anchor a child's thought processes in time	My birthday party, summer, the netball season
Sequence	Words that describe the order in which two defined events occurred	Then, before, between, first, later, next
Time	Vocabulary that picks out a specific or general moment in time	Tuesday, early March, six months ago, next year

Accuracy. Each piece of temporal information that was identified and coded for amount would then be coded for accuracy. Three measures of accuracy were used—Correct, Incorrect, and Unverifiable. The research assistant assessed each piece of temporal information and determined whether it was able to be checked against records and verified. If the child's utterance was able to be verified, it was given a 'Correct' or 'Incorrect' based on both the records and the rules set out in the coding manual, examples of which are seen in

Table 8. The full coding manual can be found in Appendix B. Details that were unable to be verified were those too vague or subjective, or where no records were available to check the accuracy of the utterance against. For example, if a child said that lockdown lasted for seven weeks, this could be verified against the dates listed on the New Zealand government's COVID-19 response website and would be determined as 'Correct'. Alternatively, if a child said that that they went for walks all the time during lockdown or that their cousin is 8 years old, these utterances were marked as 'Unverifiable' due to the respective vagueness and lack of available records.

Table 8

Excerpt of the Coding Manual for Accuracy

Question	Correct	Incorrect	Unverifiable
"How long ago did COVID-19 start?"	Response falls anytime between December 2019 and March 2020	Any other answer	"A long time ago", or similar
"When did the COVID-19 changes start for you?"	N/A	N/A	All answers due to the subjectivity of the question
"When did lockdown start?"	Response falls within 3 days either side of March 26th	Any other answer	"A long time ago", or similar
"How long did lockdown last for?"	Duration falls within 3 days either side of 49 days. "7 weeks" is also correct	Any other answer, including "2 months"	"A really long time", or similar

Inter-rater Reliability.

Time 1. The inter-rater reliability for amount and accuracy was established between two research assistants. A research assistant trained with 6 transcripts, before reliability

coding was conducted with 24 (25%) further transcripts. 'Substantial agreement' was found between the two coders (κ = .75) for amount, and 'almost perfect agreement' was found between the two coders (κ = .94) for accuracy (Viera & Garrett, 2005). Across all codes, 'almost perfect agreement' was found between the two coders (κ = .85) (Viera & Garrett, 2005).

Time 2. The inter-rater reliability for amount and accuracy was established between two research assistants. A research assistant trained with 1 transcript, before reliability coding was conducted with 12 (25.53%) further transcripts. 'Substantial agreement' was found between the two coders (κ = .75) for amount, and 'almost perfect agreement' was found between the two coders (κ = .95) for accuracy (Viera & Garrett, 2005). Across all codes, 'almost perfect agreement' was found between the two coders (κ = .85) (Viera & Garrett, 2005).

Results

The Place of Unverifiable Information

Our initial analyses included all information that was coded (both verifiable and unverifiable). However, it was observed that almost half of the information provided by children was unverifiable and could not be evaluated for its accuracy. Rates of unverifiable data were particularly high for information falling under the Sequence and Frequency codes. We therefore conducted our analyses with unverifiable information removed from the totals, to avoid under- or over-estimating children's accuracy. Analyses with the complete dataset, where the patterns of findings differed from those without unverifiable information included, are presented in Appendix C.

Data Preparation

Preliminary inspection of the final dataset revealed outliers, along with instances of severe skew and kurtosis. These were corrected through a process of winsorization, wherein

the extreme values are transformed in order to reduce the effects that they may have on the overall analyses. We did this by first identifying the outliers, and then assigning them new values that still maintained the order of the raw data. For example, an outlier that was the largest value in the dataset was winsorized to reduce the magnitude to remain the largest value, yet not as extreme. This strategy meant we could retain the data from all participants whilst protecting our results against some of the effects of outliers. The data presented in the tables throughout this section is raw and untouched—winsorized values were only used for analysis. Similarly, data was arcsine transformed for analysis of proportional measures such as accuracy.

Time 1

Analysis of temporal information reported at Time 1 was not planned for in the original interview schedule, and so only a small number of interview questions from Time 1 requested or prompted for temporal information from the children.

Total Amount of Information

As seen in Table 9, children reported very little temporal information overall (M = 6.81, SD = 2.26). The most commonly provided type of temporal information was Birthday (M = 2.22, SD = .76). The maximum score for this category was 3 which means that children were only partially reporting birthday related information. They did this most typically by reporting their date and month of birth, but not the year. Children were least likely to report Frequency information (M = 0, SD = 0), although this was because no child provided any Frequency information that was verifiable.

A simple linear regression was conducted to predict the amount of temporal information provided by children based on their age. A significant regression equation was found ($F(1,94) = 22.34$, $p < .001$), with an R^2 of .192 meaning that 19.2% of the variation in the amount given by the children can be explained by age. Children's predicted amount is

equal to 2.692+.036 (age) pieces of information when age is measured in months. Children's amount of temporal information increased by .036 for each month of age. This means that every 12 month (or one year) increase in age will result in a further .43 of a unit of information. Although this finding is significant, the practical implications are weak given that it would take more than a two-year increase in age to achieve one additional detail from the children.

Table 9

Amount of Information Reported in Time 1

	Minimum	Maximum	Mean	Std. Deviation
Time 1	2	14	6.81	2.26
Age	1	3	2.04	.35
Birthday	0	4	2.22	.76
Duration	0	3	.78	.64
Frequency	0	0	.00	.00
Landmark	0	2	.14	.37
Sequence	0	4	.10	.49
Time	0	6	1.53	1.33

Total Accuracy of Information

As seen in Table 10, children had reasonable levels of accuracy overall, yet this varied considerably within each category of temporal information. The overall mean was .76 (*SD* = .17), with scores ranging from .36 to 1. Children were most accurate when providing information about their Age (*M* = .93, *SD* = .18) and Birthday (*M* = .92, *SD* = .26). Other mean accuracy scores varied from .35 for Time (*SD* = .43) to 0 for Frequency (*SD* = 0). Table 10 shows that the standard deviations were higher than the mean for several variables (e.g., Duration), which indicates wide variability between the children's scores among the different types of temporal information.

A simple linear regression was conducted to predict the accuracy of temporal information provided by children based on their age. A significant regression equation was not found ($F(1,94) = .22$, $p = .642$), with an R^2 of .002 meaning that .2% of the variation in the children's accuracy can be explained by age. Children's predicted accuracy is equal to 2.12+.001 (age) when age is measured in months. Children's accuracy of temporal information increased by .001 for each month of age.

Table 10

Accuracy of Information Reported in Time 1

	Minimum	Maximum	Mean	Std. Deviation
Time 1	.33	1.00	.76	.17
Age	.00	1.00	.93	.18
Birthday	.00	1.00	.92	.26
Duration	.00	1.00	.10	.28
Frequency	.00	.00	.00	.00
Landmark	.00	1.00	.11	.31
Sequence	.00	1.00	.05	.22
Time	.00	1.00	.35	.43

Spontaneous Information

Children spontaneously provided very little information ($M = .81$, $SD = 1.55$). Table 11 shows means for different types of information all fell under .14 with the exception of Time ($M = .33$, $SD = .82$). The standard deviations for each type of information as well as overall were all larger than the means, which suggests that the majority of children provided very little information in each category with just a few children reporting slightly more.

Table 11

Amount of Information Reported Spontaneously in Time 1

	Minimum	Maximum	Mean	Std. Deviation
Time 1	0	9	.81	1.55
Age	0	1	.06	.24
Birthday	0	2	.06	.32
Duration	0	2	.11	.38
Frequency	0	0	.00	.00
Landmark	0	2	.14	.37
Sequence	0	4	.10	.49
Time	0	5	.33	.82

Children accurately reported just over one quarter of the information that they provided spontaneously (M = .26, SD = .42), with scores ranging from 0 to 1.

Prompted Information

Children provided more temporal information when explicitly prompted for it (M = 6, SD = 1.67). Table 12 shows they most often reported Birthday (M = 2.16, SD = .83) and Age information (M = 1.98, SD = .25), likely reflecting that they were asked scripted questions about such detail. The means, standard deviations, and ranges for Frequency, Landmark, and Sequence were all 0, presumably because there were no scripted questions prompting for these particular types of temporal information.

A simple linear regression was conducted to predict the amount of temporal information provided by children when prompted based on their age. A significant regression equation was found ($F(1,94) = 28.96$, $p < .001$), with an R^2 of .235 meaning that 23.5% of the variation in the amount given by the children can be explained by age. Children's predicted amount is equal to 2.52+.031 (age) pieces of information when age is measured in months. Children's amount of prompted temporal information increased by .031 for each month of age. This means that every 12 month (or one year) increase in age will result in a further .37

of a unit of information, meaning that again the practical significant of the association is limited.

Table 12

Amount of Information Reported when Prompted in Time 1

	Minimum	Maximum	Mean	Std. Deviation
Overall	2	10	6.00	1.67
Age	1	3	1.98	.25
Birthday	0	3	2.16	.83
Duration	0	2	.67	.54
Frequency	0	0	.00	.00
Landmark	0	0	.00	.00
Sequence	0	0	.00	.00
Time	0	4	1.20	.99

Children accurately reported just over three quarters of the prompted information (M = .77, SD = .17), with scores ranging from .33 to 1.

Time 2

In the Time 2 interviews, the interview schedule included a greater number of questions that specifically requested temporal information.

Total Amount of Information

Table 13 shows that children appeared to report much more total temporal information in the second interview than the first (M = 50.45, SD = 7.79). The most common type of temporal information that children provided was Time, followed by Landmark. The least commonly provided type of information was Frequency (M = .09, SD = .28).

A simple linear regression was conducted to predict the amount of temporal information provided by children based on their age. A significant regression equation was found ($F(1,45) = 8.86$, $p = .005$), with an R^2 of .164 meaning that 16.4% of the variation in the amount given by the children can be explained by age. Children's predicted amount is

equal to 35.942+.119 (age) pieces of information when age is measured in months. Children's amount of temporal information increased by .119 for each month of age. This means that every 12 month (or one year) increase in age will result in a further 1.43 units of information.

Table 13

Amount of Information Reported in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Overall	33	70	50.45	7.79
Age	2	3	2.23	.43
Birthday	2	3	2.77	.43
Duration	3	12	6.62	1.81
Frequency	0	1	.09	.28
Landmark	7	14	12.11	1.48
Sequence	6	14	7.60	1.74
Time	10	39	19.04	5.63

Total Accuracy of Information

Children's overall accuracy when reporting temporal information was low (M = .44, SD = .1). Table 14 shows that accuracy varied considerably across the different categories of temporal information: Children were highly accurate when talking about Age (M = .83, SD = .22) and Birthday (M = .96, SD = .12), but less so for Landmark responses (M = .57, SD = .14).

A simple linear regression was conducted to predict the accuracy of temporal information provided by children based on their age. A significant regression equation was found ($F(1,45) = 8.58$, $p = .005$), with an R^2 of .16 meaning that 16% of the variation in children's accuracy can be explained by age. Children's predicted accuracy is equal to 1.062+.003 (age) when age is measured in months. Children's accuracy of temporal

information increased by .003 for each month of age. This means that for every 12 month (or one year) increase in age children are .036 or 3.6% more accurate.

Table 14

Accuracy of Information Reported in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Overall	.20	.69	.44	.10
Age	.50	1.00	.83	.22
Birthday	.50	1.00	.96	.12
Duration	.00	.78	.18	.18
Frequency	.00	.00	.00	.00
Landmark	.23	.86	.57	.14
Sequence	.00	1.00	.39	.32
Time	.00	.65	.33	.16

Spontaneous Information

Similar to Time 1, children did not spontaneously report a great deal of temporal information ($M = 6.3$, $SD = 4.86$). Table 15 shows children most commonly talked spontaneously about Time ($M = 3.55$, $SD = 2.98$) followed by Sequence ($M = 1.04$, $SD = 1.37$). Children never spoke spontaneously about Birthday ($M = 0$, $SD = 0$), and very rarely did so about Frequency ($M = .09$, $SD = .28$) or Age ($M = .23$, $SD = .43$).

A simple linear regression was conducted to predict the amount of temporal information provided by children based on their age. Although close, a significant regression equation was not found ($F(1,45) = 3.75$, $p = .059$), with an R^2 of .077 meaning that 7.7% of the variation in the amount given by the children spontaneously can be explained by age. Children's predicted amount is equal to .107+.051 (age) pieces of spontaneous information when age is measured in months. Children's amount of temporal information increased by .051 for each month of age.

Table 15

Amount of Information Reported Spontaneously in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Time 2	0	19	6.30	4.86
Age	0	1	.23	.43
Birthday	0	0	.00	.00
Duration	0	5	.96	1.25
Frequency	0	1	.09	.28
Landmark	0	2	.43	.62
Sequence	0	5	1.04	1.37
Time	0	11	3.55	2.98

Just over half of what children reported was accurate (M = .54, SD = .31). As can be seen in Table 16, children's accuracy varied considerably both within and across categories. They were most accurate when reporting information about Time (M = .45, SD = .36).

A simple linear regression was conducted to predict the accuracy of temporal information provided spontaneously by children based on their age. A significant regression equation was found ($F(1,45) = 10.56$, $p = .002$), with an R^2 of .19 meaning that 19% of the variation in children's accuracy can be explained by age. Children's predicted accuracy is equal to -.197+.015 (age) when age is measured in months. Children's accuracy of spontaneous temporal information increased by .015 for each month of age. This means that for every 12 month (or one year) increase in age children are .18 or 18% more accurate.

Table 16

Accuracy of Information Reported Spontaneously in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Time 2	.00	1.00	.54	.31
Age	.00	1.00	.19	.40
Birthday	.00	.00	.00	.00
Duration	.00	1.00	.20	.36
Frequency	.00	.00	.00	.00
Landmark	.00	1.00	.33	.47
Sequence	.00	1.00	.33	.45
Time	.00	1.00	.45	.36

Prompted Information

When children were prompted to provide temporal information, they reported a substantial number of details (M = 44.15, SD = 4.67). Table 17 shows that the most common type of information was Time (M = 15.49, SD = 3.67), followed by Landmark (M = 11.68, SD = 1.37).

A simple linear regression was conducted to predict the amount of prompted temporal information provided by children based on their age. A significant regression equation was found ($F(1,45) = 8.401$, $p = .006$), with an R^2 of .157 meaning that 15.7% of the variation in the amount given by the children can be explained by age. Children's predicted amount is equal to 36.24+.065 (age) pieces of information when age is measured in months. Children's amount of prompted temporal information increased by .065 for each month of age. This means that every 12 month (or one year) increase in age will result in a further .78 pieces of information.

Table 17

Amount of Information Reported when Prompted in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Time 2	32	57	44.15	4.67
Age	2	2	2.00	.000
Birthday	2	3	2.77	.43
Duration	3	10	5.66	1.13
Frequency	0	0	.00	.00
Landmark	7	14	11.68	1.37
Sequence	6	9	6.55	.97
Time	9	28	15.49	3.67

Less than half of the temporal information that children reported when prompted by the interviewer were accurate (M = .42, SD = .1). Table 18 shows that, once again, children varied considerably in their accuracy both within and across categories of temporal information. Children were most accurate when talking about Birthday (M = .96, SD = .12), followed by Age (M = .83, SD = .24). They were least accurate category when reporting information about Duration (M = .14, SD = .16).

A simple linear regression was conducted to predict the accuracy of prompted temporal information provided by children based on their age. A significant regression equation was found ($F(1,45) = 5.36$, $p = .025$), with an R^2 of .106 meaning that 10.6% of the variation in children's accuracy can be explained by age. Children's predicted accuracy is equal to 1.093+.003 (age) when age is measured in months. Children's accuracy of prompted temporal information increased by .003 for each month of age. This means that for every 12 month (or one year) increase in age children are .036 or 3.6% more accurate.

Table 18

Accuracy of Information Reported when Prompted in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Time 2	.20	.67	.42	.10
Age	.50	1.00	.83	.24
Birthday	.50	1.00	.96	.12
Duration	.00	.71	.14	.16
Frequency	.00	.00	.00	.00
Landmark	.23	.86	.56	.14
Sequence	.00	1.00	.35	.35
Time	.00	.56	.29	.15

Spontaneous versus Prompted

At each time point we compared the amount of information reported spontaneously with that reported in response to interviewer prompts with a two sample t-test. At both time points, children provided significantly more information when prompted than when speaking spontaneously (Time 1: $t(95) = -29.31$, $p < .001$; Time 2: $t(46) = -48.16$, $p < .001$; see Tables 11, 12, 15, and 17 for means).

We also compared accuracy across spontaneous and prompted information at each time point. At Time 1, children were significantly more accurate when prompted than when they spontaneously reported information ($t(95) = -10.37$, $p < .001$). Conversely, at Time 2 there was no significant difference between accuracy for spontaneous compared to prompted information ($t(46) = 1.61$, $p = .114$). See discussion after Tables 11 and 12 for Time 1 means, and Tables 16 and 18 for Time 2 means.

Changes across Time

Five questions were common to both interviews and so we compared the amount and accuracy of children's responses to them (combined across the 5 questions). Children reported more information during the Time 2 interview than in the Time 1 interview ($t(46) =$

-4.25, $p < .001$: $M_{(Time\ 1)} = 6.7$, $SD = 2.3$; $M_{(Time\ 2)} = 7.83$, $SD = 1.72$). The accuracy of children's responses to the questions did not differ across interviews ($t(46) = 1.32$, $p = .097$: $M_{(Time\ 1)} = .79$, $SD = .15$; $M_{(Time\ 2)} = .75$, $SD = .14$).

Table 19

Amount and Accuracy of Responses to Five Questions asked at both Time 1 and Time 2

		Minimum	Maximum	Mean	Std. Deviation
Amount	Time 1	2	14	6.70	2.30
	Time 2	4	11	7.83	1.72
Accuracy	Time 1	.40	1.00	.79	.15
	Time 2	.43	1.00	.75	.14

Discussion

The purpose of this study was to explore how much temporal information children report when interviewed, and to evaluate the accuracy of it. We wanted to identify children's capabilities to understand how they may perform in the context of a forensic investigation and consider how that might impact the outcomes of criminal trials. Our findings will be discussed below, alongside the implications for interviewing practice and directions for future research.

Overall, we saw that children reported minimal temporal information when they were interviewed soon after the event but provided much more when more time had passed. The accuracy of the given information was often unable to be evaluated, but when it could be we saw that it varied according to the type of temporal detail the child was talking about. Even at best, children struggled to provide accurate temporal information, with the exception of knowing when their birthday is. We also examined whether children were more capable of talking about temporal information when asked directly about it rather than allowing them to provide it spontaneously. We found when specifically prompted, children provided significantly more temporal details, but the accuracy of the information was poor.

Furthermore, we generally saw that older children reported more temporal information and were more accurate than younger children.

Before we discuss our findings at length, it is worth reminding the reader that in our second interview we included scripted questions that asked children about temporal information. These questions were not part of the first interview. Our results should be interpreted in the context of an opportunistic research design which was initially focused on simply exploring children's emerging understanding of COVID-19 in their lives. Our decision to re-interview our participants, combined with the advantage of time, gave us the opportunity to reflect about the kinds of information that we were interested in learning more about which resulted in the additional temporal questions in the Time 2 interview schedule. This limited our capacity to evaluate some aspects at Time 1 as thoroughly as we could at Time 2, such as responses to prompted questions and making direct comparisons between both timepoints. This means that we cannot be certain that what we saw in Time 1 reflected a genuine lack of ability to talk about time or whether we had just not provided the optimal situation to observe this. Therefore, care is required when interpreting the data from Time 1 to avoid misrepresenting the wider implications of our findings.

How Much Temporal Information do Children Provide?

Firstly, we examined the amount of temporal detail that children provided during their interviews. We observed that children said very little about temporal features of their COVID-19 experiences at Time 1 and said much more at Time 2 (increasing from seven to fifty unique pieces of temporal information on average). This steep increase across time is inconsistent with existing research that has shown that over time children's memories fade and they recall less details about events (Hudson & Fivush, 1991; La Rooy et al., 2005; Pipe et al., 1999; Salmon & Pipe, 1997). Here the methodological difference in the addition of temporal questions to Time 2 is clear, as these questions provided not only more opportunity

to speak during a much longer interview, but required the children to respond with temporal information. In order to make a fair comparison, we matched the responses to the five temporal questions that appeared in both interviews and found that children still produced more temporal information at Time 2. This seems surprising, however it may be a consequence of the methodology used. The interviewer in Time 2 consistently asked for the child's year of birth, whereas the interviewers in Time 1 often did not and instead were satisfied if the child only responded with the date and month they were born. This means that children were much more likely to provide more information to this question at Time 2 than at Time 1, because they were given slightly more opportunity to do so. The very small amount of information provided to these five questions means that this one extra piece of information made a relatively large contribution to the data and may have skewed our findings. Alternatively, this finding may support a conclusion that children provide more detail as they increase in age, which will be discussed further below; or perhaps that when asked to discuss an event that occurred long in the past, children will automatically use temporal language. They may do this because this type of detail may be more relevant and cognitively available when an event contains a lot of temporal information, which events that happened a long time ago inherently possess. For a forensic interview, this explanation might mean interviewers need to be cautious when asking children about an event long in the past. Children may provide a lot of temporal detail, not because they know it and believe it to be correct, but because they need some temporal language in order to explain themselves better and establish their story in time. Regardless, our findings do not provide evidence that children are saying less over time due to forgetting details as was hypothesised.

Our findings show that very little of the temporal information provided by the children at either interview was provided spontaneously, and our analysis confirmed that children said significantly more when prompted by the interviewer. This is consistent with

findings that suggest children rarely offer temporal information spontaneously (Friedman & Lyon, 2005; Orbach & Lamb, 2007; Price & Evans, 2021). Children are typically socialised to guess when they are not sure of an answer, and not to challenge adults (Zajac & Brown, 2018). No child refused to answer any of the questions in the current study, and so some of their responses may have been driven by acquiescence or a desire to be compliant rather than performing a genuine memory search. This has dangerous implications in a criminal investigation context – if children endeavour to please the person conducting the interview and respond to questions without being certain about what they are saying, the accuracy of their testimony is likely to be compromised, along with their credibility.

However, Phillips and colleagues (2012) claimed that more information relevant to an investigation (e.g., people, locations, references to time) is elicited from open rather than closed questions. Our findings appear to contradict this conclusion, but when considering the types of questions that we used in our prompted section it becomes clear that we may be more consistent with Phillips and colleagues (2012) than initially thought. One of the challenges when navigating the literature in this field is that different researchers have adopted different thresholds for what is considered freely recalled information and prompted information, and within prompted information there are further differences between whether recall or recognition processes were used to answer the question. True closed questions, also known as option posing prompts, are those using recognition processes (e.g., yes or no), multiple choice, and leading questions (Brown et al., 2013). Most of the prompted questions asked in the current study are considered open-ended questions, even though we defined them as prompted. This is because they were recall-based questions which means they required the children to generate information, as opposed to confirm an option as is seen in recognition-based questions. As mentioned, for the purposes of the current study, they were considered prompted questions because they were designed to prompt a specific type of information –

temporal. With this in mind, our results are not necessarily inconsistent with Phillips and colleagues (2012). When we asked prompted questions that were nonetheless recall-based and therefore still open-ended questions, children provided information but we then began to see issues with their accuracy which perhaps was due to the more narrowed focus of the questioning. This idea of accuracy decreasing as the question narrows is well researched (Brown et al., 2013; Lamb et al., 2018). The questioning strategy adopted by the current study was consistent with forensic protocols and evidence-based recommendations. This means that recall-based processes were prioritised before asking more focussed questions when needed to elicit investigation-relevant temporal information (Lamb et al., 2018).

There was high variability in the amount of detail provided across the numerous categories of temporal information. This reflects the availability of such information to be reported. Some categories were limited by the quantity of specific questions targeting responses for that code type, the expected amount of responses generated by each question, and whether the responses were able to be verified. For example, only two questions generated prompted Age codes ("how old are you?"; "how old were you when COVID-19 started?"), and each was expected to elicit a single response. Similarly, only one question prompted Birthday codes ("when is your birthday?") and generated between one and three responses. Age and Birthday codes elicited spontaneously were often unable to be verified and thus excluded from analysis, resulting in low totals. Responses of this kind were typically the ages and birthdays of friends and family members, which the research team did not have access to information in order to verify these details. No questions prompted Frequency responses, and as a highly subjective type of temporal information any spontaneous reports of it were typically unverifiable and excluded. Therefore, it is not surprising that no children provided verifiable Frequency responses. However, the lack of Frequency responses does not provide evidence that children are inaccurate when they do talk about this type of

information, it simply means that we cannot definitively say that they were correct or incorrect. Future research might consider using staged events with the addition of prompted questions that elicit Frequency responses, to allow for analysis of the amount and accuracy of this particular type of temporal information. Questions regarding COVID-19 that would be able to be verified may include "how many times did lockdown happen?" and "how often did Prime Minister Jacinda Ardern and Dr Ashley Bloomfield come on the TV to give New Zealand an update?".

Our findings show that older children provided more pieces of temporal information than younger children, consistent with studies that have found the level of detail provided by children increases with age (Lamb et al., 2000; Orbach & Lamb, 2007; Phillips et al., 2012). This may be at least partially explained by the increase in children's vocabulary over time (Anglin et al., 1993), and perhaps the improved ability to form a narrative (Fivush, 2011; Miragoli et al., 2017). This finding shows the development of children's abilities to report temporal detail and has implications for the courtroom. Only one study to date has investigated what expectations jurors hold regarding the abilities of children to provide temporal information. Cleveland and Quas (2016) had mock jurors read fictional transcripts from a child sexual assault trial and asked them to rate how credible and accurate the child's testimony was. Participants rated children that answered temporal questions with certainty as more credible than those who were unsure, and rated younger children who were sure as more accurate than older children (Cleveland & Quas, 2016). The expectations adults report in this study seem to be in conflict with the pattern of results we've observed in current study. For this reason, it would be advantageous to the field for further research to be done so that we can deepen our understanding and identify key differences. Any misconceptions can be addressed in the form of education for the court participants which may lead them to re-evaluate their expectations of children of varying ages when they take the witness stand.

How Accurate are Children when Providing Temporal Information?

We also wanted to know how accurate children were when providing temporal detail in an interview. We observed that children were reasonably accurate at Time 1 (76%) and much less so at Time 2 (44%). At first glance, this is in alignment with research describing that the accuracy of children's recall decreases after a long delay (Hudson & Fivush, 1991; La Rooy et al., 2005; Pipe et al., 1999; Salmon & Pipe, 1997). However, it is again clear that adding new temporal questions to Time 2 impacted the results. While these prompted questions provided more opportunity for children to talk about temporal information, they also provided more opportunity for them to be incorrect when doing so. By asking children to provide this information in the absence of strong memory traces for it, we may have artificially decreased the accuracy of their responses through the compliance or acquiescence mechanisms discussed earlier. As such, interviewers should be cautious about prompting for such information as children may indeed provide it, but the information is unlikely to be reliable. When we analysed only responses to the five temporal questions that appear in both interviews, we found no significant difference in children's accuracy across time. This finding also does not align with the literature, however it may be explained by the methodology used due to three of the five matched questions yielding consistently high accuracy scores. The three questions ask for their birthday, current age, and age they were during COVID-19, which typically generated 2-3 Birthday codes and 2 Age codes. These two types of temporal information have routinely appeared as the top two most accurate categories throughout our analyses and made up a large proportion of the mean amount of detail provided by the children, which perhaps has skewed the true accuracy. Research suggests that children are highly competent when asked to identify their age (Bauer et al., 2007) and birthday (Wandrey et al., 2012), which may be explained by ages and birthdays being highly rehearsed in everyday life and therefore quite salient concepts for children to

understand and recall. Thus, when children are asked about these temporal concepts such as these that are likely to be understood and available, we may expect their responses to be more reliable than less salient details such as the precise length of COVID-19 lockdown. Further research is needed to understand when and how children might be more or less accurate. Similarly, the overall accuracy of responses of other temporal categories such as Sequence and Duration were very low. This may also be attributed to how often children are rehearsing these details and therefore the availability of this memory. It is unlikely that children had been asked to sequence the alert levels of lockdown nor to estimate how long lockdown lasted for before the interviewer in the current study requested this of them. This range in ability between categories due to the differences in use in everyday life may mean that the overall accuracy score is inflated by consistently high performing categories like Age or Birthday, and that caution should be taken when interpreting the findings.

Our findings show that children were significantly more accurate when prompted than when talking spontaneously at Time 1, but that there was no difference in accuracy between these phases at Time 2. At Time 2 the accuracy for prompted responses was lower (42%) than that of spontaneous responses (54%), although both are relatively low scores. These results are interesting when considering that responses to open-ended questions are highly accurate (Lamb et al., 2018) and tend to be more-so than those from closed questions (Lamb et al., 2007; Lamb & Fauchier, 2001). This discrepancy between findings may be explained by how children in the current study rarely offered temporal information spontaneously, and as such the greater number of prompted responses resulted in ample opportunity for accuracy verification. This means that as the amount of temporal information became greater, the opportunity for error also increased. Brown and Lamb (2015) commented that prompted questions run the risk of eliciting inaccurate information, and that seems to be the case at Time 2 as there were much more prompted questions yet no difference in accuracy between

the two types of responses. This may be explained by the fact discussed earlier that our prompted questions were recall-based, which means that accuracy is not necessarily compromised. Although we did not see any improvement in accuracy using prompted questions in Time 2, it was essential that they were asked as the children did not respond with sufficient temporal detail when speaking spontaneously. In a forensic context, interviewers will follow up open-ended questions with prompted ones to obtain any missing information required for the investigation (Lamb et al., 2007). As discussed earlier, our findings are likely to at least partially reflect children's attempts to be compliant or respond well for the interviewer. They may also reflect children's tendency to guess as well as the availability of that information – for example, a child that is asked a temporal question might recognise that they need to respond with temporal language and quickly retrieve a random temporal word. As such, the answers that children are providing cannot reliably be assumed to have been a product of them genuinely searching their memory and retrieving the appropriate information.

In line with our hypothesis, our findings indicate that at Time 2, older children provided temporal information with more accuracy than younger children. An age-related increase in accuracy was not found in Time 1. This may be due to the small amount of temporal information available to be verified for accuracy in these interviews. However, the Time 2 finding is consistent with studies that found older children performed better than younger children when asked about temporal information (Droit-Volet, 2003; Moore et al., 2014; Wandrey et al., 2012). An increase in ability may be explained by greater knowledge and understanding of time that is acquired as children age (Friedman et al., 2011). This makes sense, as expecting a child with little to no knowledge of time to talk accurately about time is a seemingly impossible task. It is an important developmental finding that children's accuracy increases as they age and may mean that children will become more reliable

witnesses in the courtroom as they progress through childhood. Although our finding was significant, age explained only 16.4% of variance of the data. Clearly the task of recalling an event after a considerable delay involves a range of abilities and is likely to reflect a broad range of factors relating to the event itself, the child, and how their memories are elicited (Lamb et al., 2018).

Sometimes the children would respond with outlandish or impossible statements that were quickly recognisable as incorrect, which raised some possibilities regarding the processes at play when children were answering the interviewer's questions. Children's poor performance either reflected that they never encoded the information as it happened and therefore couldn't recall it, that they encoded it but then found it difficult to access it later, or that they did not comprehend temporal constructs and their meaning. One such example is as seen in the title of the current study where a child stated that "lockdown was 100 months ago". This places the child's recollection of lockdown in early 2013 – a far cry from the correct answer of April 2020. Here, the child may have recognised that they were being asked a temporal question, and that they needed to respond with a temporal answer. Perhaps their thought process was that 100 is a big number, and that lockdown was a long time ago, so an answer of 100 months appears an appropriate answer to the question. Similar responses were seen from two individual children who described their most recent birthdays happening nineteen and twenty months prior to the date of the interview. These were clearly incorrect and indicate that the children lacked understanding of how often their birthday occurs, or that the length of a calendar year is twelve months. Responses such as these suggest that children are guessing, and that they either do not understand the task or never had the knowledge to answer the question correctly. This is perilous because during an investigation there is often little evidence, such as the true dates of lockdown, that can be used to verify the accuracy of a child's testimony (Phillips et al., 2012). If a child says that an occasion of sexual abuse

happened to them 100 months ago, it is then up to the judges, lawyers, and members of the jury to decide whether that is a truthful statement or not. Answers that appear outlandish could in fact still be accurate, and the consequences of withdrawing charges from a defendant for this reason may be devastating.

Future Research

Looking to the future, it would be beneficial to repeat the interview again this year. Creating a Time 3 dataset by replicating the Time 2 interview with the same participants is an excellent opportunity to make direct contrasts with the entirety of the interview questions. This would correct the key limitation of the current study which was the difference in methodology between the two interviews, and result in a rich examination of children's level of detail and accuracy across time. It would be of value to continue to repeat the interview each year as a longitudinal study to further observe the effect of longer delays on memory, however with the participant attrition rate seen during recruitment at Time 2 it seems unlikely there would be sufficient longevity within this sample. If repeating the study afresh with a new sample, it would be preferable that the participants be of a population who have had experience in a forensic setting in order to produce context-relevant and applicable results, as seen in Wandrey and colleagues (2012). It would also be interesting to conduct this interview with a sample of adults. In doing so, we would be able to evaluate whether it is only children that struggle to accurately discuss temporal information or whether it is a difficulty seen across all ages. The involvement of adults representing a wide range of ages would enable greater understanding of the development of temporal abilities across the lifespan. Moreover, this may provide a clearer understanding of at what age it is developmentally appropriate to expect people to accurately provide temporal information during a forensic investigation.

Conclusions

As seen in our results, temporal information is not something that children are likely to think about providing of their own volition. Our findings suggest that children may not encode temporal details as they experience them in everyday life or they may have difficulty in retrieving them from their memory later, and yet they may nonetheless report temporal information when asked for it. Calling back to Rosie from our hypothetical scenario, we might expect that at ten years old she would have difficulty providing temporal information but would diligently try to answer each question. Regarding her accuracy, we may expect Rosie to perform poorly which may have a large impact on her credibility—especially if her testimony is found to be inconsistent with other evidence. On the basis of these results, it appears to be developmentally inappropriate for the justice system to expect children to report temporal information accurately during an interview. Although legal practitioners may expect temporal information to be present in children's testimony of their experiences, interviewers should not seek it out in the interest of avoiding a situation where a child undermines their own credibility as well as the validity of any verdicts reached in the courtroom.

References

Akhtar, S., Justice, L. V., Knott, L., Kibowski, F., & Conway, M. A. (2018). The 'common sense' memory belief system and its implications. *The International Journal of Evidence & Proof, 22*(3), 289-304. https://doi.org/10.1177%2F1365712718784045

Anglin, J. M., Miller, G. A., & Wakefield, P. C. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development, 58*(10), 1-186. https://doi.org/10.2307/1166112

Arndt, J. (2012). False recollection: Empirical findings and their theoretical implications. *Psychology of Learning and Motivation, 56*, 81-124. https://doi.org/10.1016/B978-0-12-394393-4.00003-0

Bauer, P. J., Burch, M. M., Scholin, S. E., & Guler, 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%2Fj.1467-9280.2007.01999.x

Berman, G. L. & Cutler, B. L. (1996). Effects of inconsistencies in eyewitness testimony on mock-juror decision making. *Journal of Applied Psychology, 81*(2), 170-177. https://psycnet.apa.org/doi/10.1037/0021-9010.81.2.170

Berman, G. L., Narby, D. J., & Cutler, B. L. (1995). Effects of inconsistent eyewitness statements on mock-jurors' evaluations of the eyewitness, perceptions of defendant culpability and verdicts. *Law and Human Behavior, 19*(1), 79-88. https://doi.org/10.1007/BF01499074

Borckardt, J. J., Sprohge, E., & Nash, M. (2003). Effects on the inclusion and refutation of peripheral details on eyewitness credibility. *Journal of Applied Social Psychology, 33*(10), 2187-2197. https://doi.org/10.1111/j.1559-1816.2003.tb01880.x

Bradfield, A. L., & Wells, G. L. (2000). The perceived validity of eyewitness identification testimony: A test of the five Biggers criteria. *Law and Human Behavior, 24*(5), 581-594. https://doi.org/10.1023/A:1005523129437

Brown, D. A., & Lamb, M. E. (2015). Can children be useful witnesses? It depends how they are questioned. *Child Development Perspectives, 9*(4), 250-255. https://doi.org/10.1111/cdep.12142

Brown, D. A., Lamb, M. E., Lewis, C., Pipe, M. E., Orbach, Y., & Wolfman, Missy. (2013). The NICHD investigative protocol: An analogue study. *Journal of Experimental Psychology: Applied, 19*(4), 367-382. https://doi.org/10.1037/a0035143

Brown, D. A., Lewis, C. N., & Lamb, M. E. (2015). Preserving the past: An early interview improved delayed event memory in children with intellectual disabilities. *Child Development, 86*(4), 1031-1047. https://doi.org/10.1111/cdev.12364

Bruck, M., Ceci, S. J., & Hembrooke, H. (2002). The nature of children's true and false narratives. *Developmental Review, 22*(3), 520-554. https://doi.org/10.1016/S0273-2297(02)00006-0

Bruer, K., & Pozzulo, J. D. (2012). Influence of eyewitness age and recall error on mock juror decision-making. *Legal and Criminological Psychology, 19*(2), 332-348. https://doi.org/10.1111/lcrp.12001

Busby-Grant, J., & Suddendorf, T. (2011). Production of temporal terms by 3-, 4-, and 5-year-old children. *Early Childhood Research Quarterly, 26*(1), 87-95. https://doi.org/10.1016/j.ecresq.2010.05.002

Carter, C. A., Bottoms, B. L., & Levine, M. (1996). Linguistic and socioemotional influences on the accuracy of children's reports. *Law and Human Behavior, 20*(3), 335-358. https://doi.org/10.1007/BF01499027

Cleveland, K. C., & Quas, J. A. (2016). Adults' insensitivity to developmental changes in children's ability to report when and how many times abuse occurred. *Behavioral Sciences and the Law, 34*(1), 126-138. https://doi.org/10.1002/bsl.2211

Cossins, A., Goodman-Delahunty, J., & O'Brien, K. (2009). Uncertainty and misconceptions about child sexual abuse: Implications for the criminal justice system. *Psychiatry, Psychology and Law, 16*(3), 435-452. https://doi.org/10.1080/13218710902930234

Crimes Amendment Act 2005 (NZ). https://www.legislation.govt.nz/act/public/2005/0041/latest/DLM346175.html

Cuevas, K., Rajan, V., Morasch, K. C., & Bell, M. A. (2015). Episodic memory and future thinking during early childhood: Linking the past and future. *Developmental Psychobiology, 57*(5), 552-565. https://doi.org/10.1002/dev.21307

Droit-Volet, S. (2003). Alerting attention and time perception in children. *Journal of Experimental Child Psychology, 85*(4), 372-384. https://doi.org/10.1016/S0022-0965(03)00103-6

Estes, W. K. (1985). Memory for temporal information. In J. A. Michon & J. L. Jackson (Eds.), *Time, Mind, and Behavior* (pp. 151–168). Springer. https://doi.org/10.1007/978-3-642-70491-8_10

Fivush, R. (2011). The development of autobiographical memory. *Annual Review of Psychology, 62*(1), 559-582. https://doi.org/10.1146/annurev.psych.121208.131702

Fivush, R., Hudson, J., & Nelson, K. (1984). Children's long-term memory for a novel event: An exploratory study. *Merrill-Palmer Quarterly, 30*(3), 303-316. https://www.jstor.org/stable/23086104

Fivush, R., Kuebli, J., & Clubb, P. A. (1992). The structure of events and event representations: A developmental analysis. *Child Development, 63*(1), 188-201. https://doi.org/10.1111/j.1467-8624.1992.tb03606.x

Friedland, S. I. (1989). On common sense and the evaluation of witness credibility. *Case Western Reserve Law Review, 40*(1), 165-225. https://scholarlycommons.law.case.edu/cgi/viewcontent.cgi?article=1865&context=caselrev

Friedman, W. J. (1991). The development of children's memory for the time of past events. *Child Development, 62*(1), 139-155. https://doi.org/10.1111/j.1467-8624.1991.tb01520.x

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

Friedman, W. J. (2007). The development of temporal metamemory. *Child Development, 78*(5), 1472-1491. https://doi.org/10.1111/j.1467-8624.2007.01078.x

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., Cederborg, A. C., Hultman, E., Anghagen, O., & Magnusson, K. F. (2010). Children's memory for the duration of a paediatric consultation. *Applied Cognitive Psychology, 24*(4), 545-556. https://doi.org/10.1002/acp.1571

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

Gilbert, R., Kemp, A., Thoburn, J., Sidebotham, P., Radford, L., Glaser, D., & MacMillan, H. L. (2009). Recognising and responding to child maltreatment. *Lancet, 373*, 167–180. https://doi.org/10.1016/s0140-6736(08)61707-9

Goodman, G. S., Bottoms, B. L., Herscovici, B. B., & Shaver, R. (1989). Determinants of the child victim's perceived credibility. In S. J. Ceci, D. F. Ross, & M. P. Toglia (Eds.), *Perspectives on children's testimony* (pp. 1–22). Springer-Verlag. https://doi.org/10.1007/978-1-4613-8832-6_1

Goodman, G. S., Golding, J. M., Helgeson, V. S., Haith, M. M., & Mitchell, J. (1987). When a child takes the stand: Jurors' perceptions of children's eyewitness testimony. *Law and Human Behavior, 11*(1), 27-40. https://doi.org/10.1007/BF01044837

Goodman, G. S., Taub, E. P., Jones, D. P. H., England, P., Port, L. K., Rudy, L., Prado, L., Myers, J. E. B., & Melton, G. B. (1992). Testifying in criminal court: Emotional effects on child sexual assault victims. *Monographs of the Society for Research in Child Development, 57*(5), i-159. https://doi.org/10.2307/1166127

Goodman-Delahunty, J., Cossins, A., & O'Brien, K. (2010). Enhancing the credibility of complainants in child sexual assault trials: The effect of expert evidence and judicial directions. *Behavioral Sciences and the Law, 28*(6), 769-783. https://doi.org/10.1002/bsl.936

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

Hayne, H., & Imuta, K. (2011). Episodic memory in 3- and 4- year old children. *Developmental Psychobiology, 53*(3), 317-322. https://doi.org/10.1002/dev.20527

Henry, L., Ridley, A., Perry, J., & Crane, L. (2011). Perceived credibility and eyewitness testimony of children with intellectual disabilities. *Journal of Intellectual Disability Research, 55*(4), 385-391. https://doi.org/10.1111/j.1365-2788.2011.01383.x

Hudson, J. A., & Fivush, R. (1991). As time goes by: Sixth graders remember a kindergarten experience. *Applied Cognitive Psychology, 5*(4), 347-360. https://doi.org/10.1002/acp.2350050405

Klemfuss, J. Z., & Ceci, S. J. (2012). Legal and psychological perspectives on children's competence to testify in court. *Developmental Review, 32*(3), 268-286. https://doi.org/10.1016/j.dr.2012.06.005

La Rooy, D., Pipe, M. E., & Murray, J. E. (2005). Reminiscence and hypermnesia in children's eyewitness testimony. *Journal of Experimental Child Psychology, 90*(3), 235-254. https://doi.org/10.1016/j.jecp.2004.11.002

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

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. DOI:10.1002/9781118881248

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

Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (2000). Effects of age and delay on the amount of information provided by alleged sex abuse victims in investigative interviews. *Child Development, 71*(6), 1586-1596. https://doi.org/10.1111/1467-8624.00250

Loftus, E. F., & Bernstein, D. M. (2005). Rich false memories: The royal road to success. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 101-113). American Psychological Association. https://psycnet.apa.org/doi/10.1037/10895-008

Lum, J. A. G., Powell, M., Timms, L., & Snow, P. (2015). A meta-analysis of cross sectional studies investigating language in maltreated children. *Journal of Speech, Language, and Hearing Research, 58*(3), 961-976. https://doi.org/10.1044/2015_JSLHR-L-14-0056

Lustig, C., & Meck, W. H. (2011). Modality differences in timing and temporal memory throughout the lifespan. *Brain and Cognition, 77*(2), 298-303. https://doi.org/10.1016/j.bandc.2011.07.007

Lyon, T. D., Stolzenberg, S. N., & McWilliams, K. (2017). Wrongful acquittals of sexual abuse. *Journal of Interpersonal Violence, 32*(6), 805-825. https://doi.org/10.1177%2F0886260516657355

McCauley, M. R., & Parker, J. F. (2001). When will a child be believed? The impact of the child's age and juror gender on children's credibility and verdict in a sexual-abuse case. *Child Abuse & Neglect, 25*(4), 523–539. DOI: 10.1016/s0145-2134(01)00224-1

McCormack, T. (2015). The development of temporal cognition. In L. S. Liben, U. Muller, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Cognitive processes* (pp.624-670). John Wiley & Sons, Inc.. https://psycnet.apa.org/doi/10.1002/9781118963418.childpsy215

McCormack, T., Brown, G. D. A., Maylor, E. A., Darby, R. J., & Green, D. (1999). Developmental changes in time estimation: Comparing childhood and old age. *Developmental Psychology, 35*(4), 1143-1155. https://psycnet.apa.org/doi/10.1037/0012-1649.35.4.1143

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%2F0886260516645812

Miragoli, S., Camisasca, E., & Di Blasio, P. (2017). Narrative fragmentation in child sexual abuse: The role of age and post-traumatic stress disorder. *Child Abuse & Neglect, 73*(1), 106-114. https://doi.org/10.1016/j.chiabu.2017.09.028

Moore, B. D., Brooks, P. J., & Rabin, L. A. (2014). Comparison of diachronic thinking and event ordering in 5- to 10-year-old children. *International Journal of Behavioral Development, 38*(3), 282-292. https://doi.org/10.1177%2F0165025414520806

Morison, S., & Greene, E. (1992). Juror and expert knowledge of child sexual abuse. *Child Abuse and Neglect, 16*(4), 595-613. DOI: 10.1016/0145-2134(92)90075-3

New Zealand Police. (n.d.). *Victimisations (demographics).* Retrieved March 16, 2022, from https://www.police.govt.nz/about-us/publications-statistics/data-and-statistics/policedatanz/victimisations-demographics

Oranga Tamariki–Ministry for Children. (2021). *Care and protection – statistics.* Retrieved March 16, 2022, from https://www.orangatamariki.govt.nz/about-us/performance-and-monitoring/quarterly-report/care-and-protection-statistics/

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

Ost, J., Easton, S., Hope, L., French, C. C., & Wright, D. B. (2017). Latent variables underlying the memory beliefs of chartered clinical psychologists, hypnotherapists, and undergraduate students. *Memory, 25*(1), 57-68. https://doi.org/10.1080/09658211.2015.1125927

Otgaar, H., Candel, I., & Merckelbach, H. (2008). Children's false memories: Easier to elicit for a negative than for a neutral event. *Acta Psychologica, 128*(2), 350-354. https://doi.org/10.1016/j.actpsy.2008.03.009

Otgaar, H., Smeets, T., & Peters, Maarten. (2012). Children's implanted false memories and additional script knowledge. *Applied Cognitive Psychology, 26*(5), 709-715. https://doi.org/10.1002/acp.2849

Pathman, T., Doydum, A., & Bauer, P. J. (2013). Bringing order to life events: Memory for the temporal order of autobiographical events over an extended period in school-aged children and adults. *Journal of Experimental Child Psychology, 115*(2), 309-325. https://doi.org/10.1016/j.jecp.2013.01.011

Peterson, C. (2011). Children's memory reports over time: Getting both better and worse. *Journal of Experimental Child Psychology, 109*(3), 275-293. https://doi.org/10.1016/j.jecp.2011.01.009

Phillips, E., Oxburgh, G., Gavin, A., & Myklebust, T. (2012). Investigative interviews with victims of child sexual abuse: The relationship between question type and investigation relevant information. *Journal of Police and Criminal Psychology, 27*(1), 45-54. https://doi.org/10.1007/s11896-011-9093-z

Pipe, M. E., Gee, S., Wilson, J. C., & Egerton, J. M. (1999). Children's recall 1 or 2 years after an event. *Developmental Psychology, 35*(3), 781-789. https://psycnet.apa.org/doi/10.1037/0012-1649.35.3.781

Pipe, M. E., Sutherland, R., Webster, N., Jones, C., & La Rooy, D. (2004). Do early interviews affect children's long-term event recall? *Applied Cognitive Psychology, 18*(7), 823-839. https://doi.org/10.1002/acp.1053

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

Price, H. L., & Evans, A. D. (2021). With support, children can accurately sequence within-event components. *Applied Cognitive Psychology, 35*(4), 890-899. https://doi.org/10.1002/acp.3816

Puddifoot, K. (2020). Re-evaluating the credibility of eyewitness testimony: The misinformation effect and the overcritical juror. *Episteme, 17*(2), 255-279. https://doi.org/10.1017/epi.2018.42

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

Roberts, K. P., Brubacher, S. P., Drohan-Jennings, D., Glisic, U., Powell, M. B., & Friedman, W. J. (2015). Developmental difference in the ability to provide temporal information about repeated events. *Applied Cognitive Psychology, 29*(3), 407-417. https://doi.org/10.1002/acp.3118

Rogers, P., & Davies, M. (2007). Perceptions and perpetrators in a depicted child sexual abuse case. *Journal of Interpersonal Violence, 22*(5), 566–584. https://doi.org/10.1177/0886260506298827

Ross, D. F., Jurden, F. H., Lindsay, R. C. L., & Keeney, J. M. (2003). Applications and limitations of a two-factor model of child witness credibility. *Journal of Applied Social Psychology, 33*(2), 418–430. https://doi.org/10.1111/j.1559-1816.2003.tb01903.x

Salmon, K. (2001). Remembering and reporting by children: The influence of cues and props. *Clinical Psychology Review, 21*(2), 267-300. https://doi.org/10.1016/S0272-7358(99)00048-3

Salmon, K., & Pipe, M. E. (1997). Props and children's event reports: The impact of a 1-year delay. *Journal of Experimental Child Psychology, 65*(3), 261-292. https://doi.org/10.1006/jecp.1996.2362

Saywitz, K. J. (2002). Developmental underpinnings of children's testimony. In H. L. Westcott, G. M. Davies, & R. Bull (Eds.), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice* (pp. 3–19). John Wiley & Sons. https://doi.org/10.1002/9780470713679.ch1

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

Sharman, S. J., Powell, M. B., & Roberts, K. P. (2011). Children's ability to estimate the frequency of single and repeated events. *International Journal of Police Science and Management, 13*(3), 234-242. https://doi.org/10.1350%2Fijps.2011.13.3.243

Tabak, S. J., & Klettke, B. (2014). Mock jury attitudes towards credibility, age, and guilt in a fictional child sexual assault scenario. *Australian Journal of Psychology, 66*(1), 47-55. https://doi.org/10.1111/ajpy.12035

Tillman, K. A., & Barner, D. (2015). Learning the language of time: Children's acquisition of duration words. *Cognitive Psychology, 78*(1), 57-77. https://doi.org/10.1016/j.cogpsych.2015.03.001

Tizzard-Drover, T., & Peterson, C. (2004). The influence of an early interview on long-term recall: A comparative analysis. *Applied Cognitive Psychology, 18*(6), 727-743. https://doi.org/10.1002/acp.1028

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. Family Medicine, 37(5), 360-363. https://pubmed.ncbi.nlm.nih.gov/15883903/

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://psycnet.apa.org/doi/10.1037/a0024812

Zajac, R., & Brown, D. A. (2018). Conducting successful memory interviews with children. *Child Adolescent Social Work Journal, 35*, 297-308. https://doi.org/10.1007/s10560-017-0527-z

Appendices

Appendix A – Coding Manual (Amount)

Coding Rules

Pre-/post-amble

- Discussion that is either side of the prescribed interview schedule is not coded
- Begin coding from the assent script
- Stop coding after the interviewer thanks the child for participating

Clarity

- If it is unclear what the child is talking about, do **not** code it
- Only code what is there on the page – do **not** infer what they mean in order to code something
- Exception: when an incorrect word is used (due to forgetting, not knowing, etc) but their intention is exceptionally clear
    - E.g., "long down" used instead of "lockdown"

A single utterance may contain different phrases that belong in different categories from each other.

- E.g, "Every Thursday" is a single utterance
    - 'Every' belongs in Frequency, 'Thursday' belongs in Time

Corrections

- Information that is retracted by the child and replaced by a new piece of information
    - The new piece of information is coded, and the original code is removed

"Before"

- 'Before' is **not** considered a sequence code when in the context of "I had never done ___ before" as its function is not as a time anchor
- When 'before' is introduced by the interviewer, it is **not** coded
    - Exception: Different contexts

Repetitions

- If two or more utterances refer to the same context, do **not** code more than one
- The initial utterance should be coded, and any further should be marked as repetitions
    - If two coders have coded the same utterance referring to the same context but at different points in the transcript, this is counted as an agreement as both coders have agreed that this piece of information is codeable
        - If one coder has coded all repetitions and the other has coded just one, this is 1 agreement and X amount of disagreements

Counting and quantity

- Some utterances may appear to be time-based, but the child has just counted something, and it is not time-based
    - E.g, "We did three things"
        - 'Three', although a number, is **not** referencing time
    - E.g., 'A lot'
        - "We did a lot of school work", 'a lot' is **not** coded
            - I.e., 'heaps' (quantity)
        - "We went to the playground a lot", 'a lot' is coded
            - I.e., 'often' (Frequency)

False starts

- False starts are instances where the child begins a sentence but changes direction and never finishes the sentence
- Ignore and do **not** code

Phrases that are both activities and time-based pieces of information

- E.g., lunchtime, afternoon tea
    - Pay attention to the context and code the time aspect
    - "We had afternoon tea", 'afternoon tea' is **not** coded
    - "It happened at afternoon tea", 'afternoon tea' is coded
    - "It was afternoon tea time", 'afternoon tea' is coded

Interviewer's speech

- Nothing that the interviewer says is codeable
- Any information introduced by the interviewer and later repeated by the child is **not** codeable
    - E.g., I: "and then what happened?" C: "then we did maths", 'then' is **not** codeable
    - E.g., I: "tell me all about what a weekend was like before Covid-19" C: "well, before Covid-19…", 'before' is **not** codeable

"Finally"

- Finally is not considered a Sequence code when in the context of "I could finally do more stuff with my friends" as its function is not as a time anchor

Lockdown language

- Lockdown language such as "COVID-19", "lockdown", "Level _", etc, is **not** codeable at any point
- Sequence codes may exist in the context of lockdown language – these are codeable
    - E.g, "before lockdown", 'before' is coded as Sequence
- **Exception:** Specific responses in Section 5. See Time 2 Specific Rules for clarification

"In" "During"

- Some utterances may appear as Landmark codes e.g., "in the school holidays"
- "In" places the utterance within a time period and thus is coded as a Time code
    - "In the school holidays" = Time
    - "The school holidays" = Landmark
    - "It was in winter" = Time
    - "It was winter" = Landmark

Different contexts

- If the same utterance is used to describe a different event/context from the first occurrence, then this is considered a new unique piece of information and is coded

- E.g, a child uses "eight or nine weeks" to describe the length of time spent inside their bubble. The child then uses "eight or nine weeks" to describe how long they didn't have to go to school. These are two different contexts, thus two individual codes

"More"

- If "more" is used in a sentence that would still make sense if you turned it into "more often", then it is coded for Frequency
    - E.g., "we had to stay home more [often]"
- "More" used in other contexts is **not** coded
    - "I'll just play around at home and watch more YouTube"
        - The child is not watching YouTube more often, the child is continuing to watch YouTube

Experiences not their own

- Children describing an experience that is not directly their own such as what they saw in a video or friends going on holidays **is** codeable

Time-based nouns

- Nouns such as "day" "week" and "year" are coded in the context of the utterance, and are **not** coded individually as nouns as they are non-specific
- More descriptive/specific nouns such as "Tuesday" and "August" are coded for Time as they are specific
- Example: "every day"
    - Correct: "every day" = Frequency
    - **Incorrect:** "every" = Frequency and "day" = Time
- Example: "every Tuesday"
    - Correct: "every" = Frequency, "Tuesday" = Time

"One time" vs "Once"

- "One time" refers to a particular time when something happened, and is therefore a Time code
    - E.g., "one time we watched a video" (colloquial, similar to "one day")

- "Once" refers to something occurring a singular time, and is therefore a Frequency code
    - E.g., "once we watched a video" (they did not watch a video twice)
- Use contextual clues to come to a decision on how to code these utterances

"The last day of school"

- This is a Time code when discussing an activity or event that happened on the last/final day of school, as it describes a moment in time
- This a Sequence code when in the context of "and then it was the last/final day of school", as it completes a sequence of events
    - Use contextual clues to come to a decision on how to code these utterances

Time 2 Specific Rules

Amount coding in the Birthday Guessing Game & end of Section 5

- "Near" questions
    - Responses coded as Prompted Landmarks
        - Includes Yes/No answers
- "Before/After" questions
    - Responses coded as Prompted Sequence
- "Birthday/Easter happen during lockdown?" questions
    - Responses coded as Prompted Landmarks
        - Includes Yes/No answers
- "What Alert Level were we in Birthday/Easter" questions
    - Responses coded as Prompted Time

"Tell me the order that the Alert Levels happened in"

- Entire response to be coded as one single Prompted Sequence code
    - Any words typically coded as Sequence within this are **not** to be coded

"Was lockdown *as long as*…"

- All three given options to be coded as Prompted Landmarks

  - The summer holidays; a school term; the mid-year school holidays

"What season was it when lockdown happened?"

- All four possible answers to be coded as Prompted Landmarks

Top tips to keep in mind when coding

- Always think about the phrase's function in the sentence
- Would the sentence still make sense if you changed the target phrase to a different one from the suspected code category? How about a different code category?

Watch out for utterances such as "then" – children often use this word as a gap filler (similar to "and") and actually has no time-based function

Appendix B – Coding Manual (Accuracy)

Coding Rules

- All responses finalised as Amount codes will then be marked for Accuracy
    - Correct: Response is correct and accurate according to the below guide
    - Incorrect: Response is incorrect and inaccurate according to the below guide
    - Unverifiable: Response is unable to be marked as either Correct or Incorrect due to either not having proof in order to verify it or by being overly subjective
- Dates must be calculated from the **date of interview**, not today's date
- Difficult descriptors
    - One month is described as the same date in the previous/following calendar month
        - E.g, "two months ago" said on July 14th refers to May 14th
    - One week is described as 7 days
        - E.g, "two weeks ago" said on July 14th refers to June 30th
    - "Early", "mid", and "late"
        - In a month: Early = 1st-10th, mid = 11th-20th, late = 21st-31st
- If it is unclear what the topic of the temporal piece of information is, and the accuracy code would change depending on this, code it as Unverifiable
- Estimations such as "nearly" or "just over" are Unverifiable
- "_____ OR _____"
    - Instances where a child has given multiple options instead of a single, definitive answer are coded as Incorrect
- "Within 3 months either side of ___"
    - Refers to the child being in either the Recent or Forthcoming birthday group for that particular landmark

Length of the summer holidays, the school term, and the mid-year school holidays

- The summer holidays = 6 weeks
- The school term = 10 weeks
    - 2.5 months
- The mid-year school holidays = 2 weeks

Introductory Questions

Question	**Correct**	**Incorrect**	**Unverifiable**
Is Easter near your birthday?	Within 3 months either side of Easter in year of interview (2nd January-5th July 2021)	Outside of 3 months either side of Easter in year of interview	
Is Easter [2021] before or after your birthday?			Always
Is it near your birthday right now?	Within 3 months either side of interview	Outside of 3 months either side of interview	
Is right now before or after your birthday?			Always
Is Christmas near your birthday?	Within 3 months either side of Christmas (September 25th – March 25th)	Outside of 3 months either side of Christmas	
Is Christmas before or after your birthday?			Always
How long ago was your birthday?	Answer falls within the same month as birthday Answer falls within 1 week either side of birthday	Any other answer	
When is your birthday?	Correct date	Incorrect date	

How old are you right now?	Correct age in years	Incorrect age in years	
How old were you when COVID-19 started?	Age in December 2019, or January, February, or March 2020	Age in any other month	

Section 1–4 Questions

Question	**Correct**	**Incorrect**	**Unverifiable**
How long ago did COVID-19 start?	Answer falls anytime between December 2019 and March 2020	Any other answer	
When did the changes start for you?			All answers
How long did you have to stay in your bubble for?	Duration falls within 3 days either side of 49 days An answer of '7 weeks' is also correct	Any other answer, including '2 months'	

Section 5 Questions

Question	**Correct**	**Incorrect**	**Unverifiable**
How long ago was lockdown?	Answer falls anytime between March 26th and May 13th	Any other answer	
What season was it when lockdown happened?	Autumn	Winter, spring, summer	
When did lockdown start?	Date falls within 3 days either side of March 26th An answer of 'March' is also correct	Any other answer	
When did lockdown finish?	Date falls within 3 days either side of May 13th	Any other answer	

	An answer of 'May' is also correct		
How long did lockdown last for?	Duration falls within 3 days either side of 49 days An answer of '7 weeks' is also correct	Any other answer, including '2 months'	
Was lockdown as long as *[the summer holidays/a school term/the mid-year school holidays]*?	The summer holidays	A school term, the mid-year school holidays	
There were four levels of lockdown. Tell me what order they happened in.	**A maximum of 6 points:** 1 point for stating four alert levels, excluding the preceding 2-3 (i.e, participant does not continue to list levels) 1 point for each alert level in its correct slot (e.g, "4-2-3-1" receives 2 out of 4 points) 1 bonus point for the preceding 2-3 [-4-3-2-1]	Points that are not received as correct are incorrect E.g, 4/6 correct results in 2/6 incorrect	
When did Level 4 start?	Date falls within 3 days either side of March 26th An answer of 'March' is also correct	Any other answer	
How long did Level 4 last for?	Estimation falls within 3 days either side of 33 days An answer of '5 weeks' is also correct An answer of '1 month' is also correct	Any other answer	

Was Level 4 as long as *[the summer holidays/a school term/the mid-year school holidays]*?	The summer holidays	A school term, the mid-year school holidays	
When did Level 3 start?	Date falls within 3 days either side of April 28th An answer of 'April' is also correct	Any other answer	
How long did Level 3 last for?	Estimation falls within 3 days either side of 16 days An answer of 'two weeks' is also correct An answer of 'half a month' is also correct	Any other answer	
Was Level 3 as long as *[the summer holidays/a school term/the mid-year school holidays]*?	The mid-year school holidays	The summer holidays, a school term	
When did Level 2 start?	Date falls within 3 days either side of May 14th An answer of 'May' is also correct	Any other answer	
How long did Level 2 last for?	Estimation falls within 3 days either side of 26 days An answer of '4 weeks' is also correct An answer of '1 month' is also correct	Any other answer	
Was Level 2 as long as *[the summer holidays/a school term/the mid-year school holidays]*?	The mid-year school holidays	The summer holidays, a school term	

When did Level 1 start?	Date falls within 3 days either side of June 9th An answer of 'June' is also correct	Any other answer	
How long did Level 1 last for?	Estimation falls within 3 days either side of 64 days An answer of '9 weeks' is also correct An answer of '2 months' is also correct	Any other answer	
Was Level 1 as long as *[the summer holidays/a school term/the mid-year school holidays]*?	None	The summer holidays, a school term, the mid-year school holidays	
Did your birthday last year *[your Xth birthday]* happen during lockdown?	Birthday occurred between March 26th and May 13th	Birthday did not occur between March 26th and May 13th	
Yes: What Alert Level was New Zealand in when your birthday happened?	Correct Alert Level is identified	Incorrect Alert Level is identified	
No: Was lockdown near your birthday?	Within 3 months either side of lockdown	Outside of 3 months either side of lockdown	
No: Was lockdown before or after your birthday?			Always
Did Easter happen during lockdown?	Yes	No	

Yes: What Alert Level was New Zealand in when Easter happened?	Level 4	Any other Alert Level	
No: Was lockdown near Easter?		All answers	
No: Was lockdown before or after Easter?		All answers	

Key dates

Easter 2020: Friday 10th – Monday 13th April

Easter 2021: Friday 2nd – Monday 5th April

Level 4: 25th March – 27th April

- 33 days

Level 3: 23rd – 25th March; 28th April – 13th May

- 2.5 days; 16 days

Level 2: 21st – 23rd March; 14th May – 8th June

- 2 days; 26 days

Level 1: 9th June – 11th August

- 64 days

Auckland: Entered Level 3 on August 12th 2020

- Rest of the country entered Level 2

First case in New Zealand: 28th February

Order of Alert Levels: 2 – 3 – 4 – 3 – 2 – 1

Duration of 'Lockdown' (L4+L3) = 7 weeks = 49 days

Summer holidays = ~6 weeks

A school term = ~10 weeks

The mid-year school holidays = ~2 weeks

Appendix C – Unverifiable Data and Analyses

This appendix displays findings and analyses with the complete dataset that includes unverifiable information. Refer to the Results section (beginning page 27) and Tables 9–19 to observe how the following findings differ from those excluding unverifiable information.

Time 1

Table 1

Amount of Information Reported in Time 1

	Minimum	Maximum	Mean	Std. Deviation
Time 1	7	178	59.61	32.90
Age	1	7	2.42	.91
Birthday	0	5	2.3	.89
Duration	0	28	5.42	5.32
Frequency	0	47	15.81	10.54
Landmark	0	9	1.24	1.72
Sequence	0	60	15.03	11.04
Time	1	68	17.38	12.11

Table 2

Accuracy of Information Reported in Time 1

	Minimum	Maximum	Mean	Std. Deviation
Time 1	.02	.57	.11	.24
Age	.00	1.00	.84	.24
Birthday	.00	1.00	.89	.27
Duration	.00	1.00	.03	.13
Frequency	.00	.00	.00	.00
Landmark	.00	1.00	.06	.20
Sequence	.00	.31	.01	.03
Time	.00	.25	.04	.06

Table 3

Amount of Information Reported Spontaneously in Time 1

	Minimum	Maximum	Mean	Std. Deviation
Time 1	2	161	51.12	31.65
Age	0	5	.39	.88
Birthday	0	3	.17	.56
Duration	0	27	4.64	4.87
Frequency	0	47	15.70	10.51
Landmark	0	8	1.16	1.61
Sequence	0	60	14.85	10.95
Time	0	59	14.23	11.29

Children accurately reported close to none of the information that they provided spontaneously (*M* = .01, *SD* = .02), with scores ranging from 0 to .14.

Table 4

Amount of Information Reported when Prompted in Time 1

	Minimum	Maximum	Mean	Std. Deviation
Overall	2	17	8.49	2.67
Age	1	4	2.03	.34
Birthday	0	3	2.16	.83
Duration	0	3	.78	.60
Frequency	0	2	.11	.35
Landmark	0	2	.08	.31
Sequence	0	2	.18	.41
Time	0	9	3.15	1.96

Children accurately reported just over half of the prompted information (*M* = .56, *SD* = .18), with scores ranging from .14 to 1.

Time 2

Table 5

Amount of Information Reported in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Overall	45	181	99.13	32.67
Age	2	4	2.36	.605
Birthday	2	9	3.09	1.12
Duration	3	29	11.45	5.46
Frequency	0	35	11.11	7.11
Landmark	7	20	13.77	2.34
Sequence	11	47	23.49	9.40
Time	11	66	33.87	13.82

Table 6

Accuracy of Information Reported in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Overall	.11	.48	.24	.07
Age	.50	1.00	.80	.22
Birthday	.22	1.00	.91	.19
Duration	.00	.70	.12	.14
Frequency	.00	.00	.00	.00
Landmark	.19	.86	.51	.13
Sequence	.00	.50	.14	.14
Time	.00	.44	.20	.10

Table 7

Amount of Information Reported Spontaneously in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Time 2	2	123	48	29.56
Age	0	2	.36	.61
Birthday	0	7	.30	1.12
Duration	0	23	5.62	5.32
Frequency	0	35	10.98	7.12
Landmark	0	8	1.74	1.81
Sequence	0	35	13.13	9.44
Time	0	45	15.87	11.44

Table 8

Accuracy of Information Reported Spontaneously in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Time 2	.00	.30	.08	.06
Age	.00	1.00	.18	.38
Birthday	.00	.00	.00	.00
Duration	.00	.67	.06	.14
Frequency	.00	.00	.00	.00
Landmark	.00	1.00	.18	.31
Sequence	.00	.30	.05	.08
Time	.00	.50	.12	.13

Table 9

Amount of Information Reported when Prompted in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Time 2	38	70	51.13	5.61
Age	2	2	2.00	.000
Birthday	2	3	2.79	.41
Duration	3	10	5.83	1.24
Frequency	0	1	.13	.34
Landmark	7	15	12.02	1.57
Sequence	8	13	10.36	1.13
Time	9	34	18.00	4.63

Table 10

Accuracy of Information Reported when Prompted in Time 2

	Minimum	Maximum	Mean	Std. Deviation
Time 2	.16	.58	.36	.09
Age	.50	1.00	.83	.24
Birthday	.50	1.00	.96	.13
Duration	.00	.71	.14	.16
Frequency	.00	.00	.00	.00
Landmark	.23	.86	.55	.14
Sequence	.00	.60	.22	.22
Time	.00	.48	.25	.13

Spontaneous versus Prompted

At Time 1, children provided significantly more information when speaking spontaneously than when prompted ($t(95) = 13.654$, $p < .001$). Conversely, at Time 2 there was no significant difference between amount for spontaneous compared to prompted information ($t(46) = -.786$, $p = .436$). In the current Appendix, see Tables 3 and 4 for Time 1 means, and Tables 7 and 9 for Time 2 means.

At both time points, children were significantly more accurate when prompted than when speaking spontaneously (Time 1: $t(95) = -33.658$, $p = <.001$; Time 2: $t(46) = -20.213$, $p = <.001$). In the current Appendix, see discussion after Tables 3 and 4 for Time 1 means, and Tables 8 and 10 for Time 2 means.

Changes across Time

Table 11

Amount and Accuracy of Responses to Five Questions asked at both Time 1 and Time 2

		Minimum	Maximum	Mean	Std. Deviation
Amount	Time 1	3	16	7.74	3.04
	Time 2	4	22	8.83	2.99
Accuracy	Time 1	.25	1.00	.71	.19
	Time 2	.23	1.00	.69	.15

Children reported more information during the Time 2 interview than in the Time 1 interview ($t(46) = -2.11$, $p < .040$). The accuracy of children's responses to the questions did not differ across interviews ($t(46) = 1.01$, $p = .160$)

Regression Analyses

The following regression analyses were conducted with the total dataset for amount and accuracy at both timepoints. As such, there are no separate analyses for spontaneous and prompted information as may be seen in the Results section.

Time 1 Amount

A significant regression equation was found ($F(1,94) = 19.45$, $p < .001$), with an R^2 of .171 meaning that 17.1% of the variation in the amount given by the children can be explained by age. Children's predicted amount is equal to 1.273+.520 (age) pieces of information when age is measured in months. Children's amount of temporal information increased by .52 for each month of age. This means that every 12 month (or one year) increase in age will result in a further 6.24 units of information.

Time 1 Accuracy

A significant regression equation was not found ($F(1,94) = 1.02$, $p = .316$), with an R^2 of .011 meaning that 1.1% of the variation in the children's accuracy can be explained by age. Children's predicted accuracy is equal to .754+-.001 (age) when age is measured in months. Children's accuracy of temporal information decreased by .001 for each month of age.

Time 2 Amount

A significant regression equation was found ($F(1,45) = 17.38$, $p < .001$), with an R^2 of .279 meaning that 27.9% of the variation in the amount given by the children can be explained by age. Children's predicted amount is equal to 19.932+.652 (age) pieces of information when age is measured in months. Children's amount of temporal information increased by .65 for each month of age. This means that every 12 month (or one year) increase in age will result in a further 7.82 units of information.

Time 2 Accuracy

A significant regression equation was not found ($F(1,94) = .038$, $p = .846$), with an R^2 of .001 meaning that .1% of the variation in the children's accuracy can be explained by age. Children's predicted accuracy is equal to 1.032+.000 (age) when age is measured in months. Children's accuracy of temporal information increased by .000 for each month of age.