

GROUND RULES IN FORENSIC INTERVIEWS: SHOULD WE MAKE PRACTICE
MORE APPLIED?

BY

KELLY JANINE LAWRENCE

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Master of Science

Victoria University of Wellington
2021

Acknowledgements

Thank you to Dr Deirdre Brown, a reliable beacon of positive reassurance, for setting me on track (multiple times) throughout the year and providing insightful and helpful feedback each time. I deeply appreciate my coding partner, Hannah Kitchen, and her endless patience during our many afterhours discussions on the intricate nuances of the way children speak about past experiences. I'm extremely grateful for my mother, who, when I was ready to click submit and forget about it, painstakingly used time she should be relaxing on holiday to give me comprehensive and very constructive feedback on the final draft of this thesis. Finally, during the unforgettable events of 2020, I thank and appreciate everyone who has helped keep my mental health in relatively good shape. This is especially the case for my partner, Arash, whose unwavering support and belief in me during the everyday has helped ground me when I felt untethered. Thank you.

Table of Contents

Abstract	4
Method	15
Participants	15
Procedure.....	15
Data Processing	22
Results.....	24
Preliminary Analysis	26
Main Analysis.....	26
Discussion	30
Does Increasing the Match Between Ground Rules Practice and Interview Context Facilitate Children’s Use of the Rules?	30
Does the Match Between Ground Rules Practice and Interview Context Affect How Accurate Children Are When Reporting a Recent Event?	34
Is Use of Ground Rules Associated with the Accuracy of Details Children Report About a Recent Event?	36
Limitations.....	36
Conclusion.....	37
References	39
Appendix A	43
Appendix B	51
Appendix C	55

Abstract

Introducing ground rules is recommended in many forensic interview best-practice protocols, but children do not always use them when they should. There is not yet a consensus in the literature on the best way to teach the rules, and many of the practice methods researched are not feasible for practitioners. Additionally, increased intensity of practice can lead to adverse effects on other aspects of child testimony too. We draw on cognitive learning literature to understand how to better facilitate ground rule use amongst children in forensic interviews. Ninety-three children between the ages of 5-12 from Greater Wellington region, New Zealand, participated in a staged event at their school and were interviewed using the National Institute of Child Health and Human Development (NICHD) Protocol (Lamb et al., 2018) 2-3 weeks later. At the interview, children practised the ground rules 'I don't know' (IDK), 'I don't understand' (IDU) and 'Correct me' (CM) in one of four ways which varied by the degree of match between the practice and interview context. Children were asked difficult questions designed to elicit the rules throughout the interview, and coding children's accuracy of reporting also examined the broader effects of practice method and rule use. No significant effects were found between the practice method and responses to difficult questions for the IDK and CM rules. The Control condition, which received no ground rules instruction or practice, was significantly different to the other practice conditions for the IDU rule. In addition to this, there was no significant effect of practice method or competency at using ground rules on children's general accuracy about the event. Several possible explanations for this pattern of results are explored in the discussion section.

Few times in a person's life do words carry such weight than when in a forensic interview. This is especially when the person being interviewed is a child and maybe the only source of information about the event(s) in question, as commonly occurs in sexual abuse investigations. For this reason, forensic interviewers must be equipped with the right skills and knowledge to support interviewees to provide accurate reports. As children's cognitive and social skills are still forming, interview approaches which work well with adults can lead to miscommunication and comprehension errors when used with children (Lamb et al., 2018). Consequently, interviewers must use particular strategies and questioning techniques adapted for use with children (Lamb et al., 2018).

One strategy commonly used around the world, including the United Kingdom, United States of America and New Zealand, are 'ground rules' (Anderson., 2013; Ministry of Justice, 2011; Westera et al., 2016). Ground rules are an endorsed shared set of rules the child can draw upon to answer questions they do not know the answer to, are unclear, incomprehensible or incorrect (Brubacher et al., 2015). Introducing these rules is believed to reduce the social pressure children may feel to provide answers to questions outside the bounds of their knowledge and comprehension and cue them to monitor their knowledge and comprehension in the process (Brubacher et al., 2015).

However, past research has reached no consensus on the best way to introduce ground rules to facilitate their use. Many of the practice methods associated with increased use of ground rules are also not feasible for practitioners to implement because they do not account for time constraints or preservation of child motivation for the more challenging substantive phase of the interview (Nesbitt & Markham, 1999; Peters & Nunez, 1999; Saywitz et al., 1999; Saywitz & Moan-Hardie, 1994). Additionally, some practice methods compromise other vital aspects of child testimony (Gee et al., 1999; Nesbitt & Markham, 1999; Saywitz & Moan-Hardie, 1994). This leads us to question the broader effects of ground rules practice and use beyond children's observable ability to resist difficult questions that should elicit the ground rules. This research project draws upon the ground rules and cognitive learning literature bases to explore a more flexible way of practising the ground rules that considers the challenges mentioned above. Specifically, we examined if increasing the match between how ground rules are practised and then applied in an interview would help facilitate children's use of them when they encounter difficult questions. In addition to this, we explored if the way children practised ground rules also affected how accurate they were in an interview following the commonly used National Institute of Child Health and Human Development (NICHD) protocol (Lamb et al.,

2018). Finally, to understand how ground rule use might more holistically impact interview quality, we assessed if increased proficiency in use of ground rules in response to difficult questions was also associated with how accurate children were when reporting about a past experienced event.

So why are ground rules important for children in forensic interviews? Firstly, introducing ground rules is believed to assist children in adapting to a forensic interview's unique context. In day-to-day life, children learn how to organise and verbalise their memories through scaffolded conversations with those around them (Nelson & Fivush, 2004). As such, they are used to being the junior partners in conversation, with adults signalling what information children should contribute (Fivush et al., 2006; Lamb & Brown, 2006). These experiences then shape children's expectations of what they should share in conversations. However, this is problematic in a forensic interview because its unique context demands that children take on an unfamiliar discursive position (Salmon & Reese, 2015). In this new position, the child is now the expert informant, and the adult is naïve (Lamb & Brown, 2006). Because this is not how children are typically positioned in exchanges with adults, they might assume the forensic interviewer already knows about the event(s) in question or might not recognise the importance of reporting key details. Introducing children to ground rules is meant to challenge a child's implicit expectations of the interview by letting them know the usual conversational rules do not apply, and at the same time, providing them with new rules to use instead.

Ground rules are also important because children often attempt to answer questions they do not understand. This effect was first reported by Hughes and Grieses (1980), who asked a group of 5 and 7-year-olds bizarre questions such as, "Is milk bigger than water?". Almost every child attempted to answer the questions, seemingly without realising they were not technically answerable (Hughes & Grieses, 1980). This effect has been replicated in larger sample sizes using different question types too (Pratt; 1990; Waterman et al., 2000). Pratt (1990) found that it did not matter whether the questions were semantically-related (e.g. "Is a fork happier than a knife?) or not (e.g. "Is a cup sadder than an orange?"); children still tried to answer them. However, when the children were asked later if the questions were 'silly' or not, they could correctly discern between answerable and unanswerable questions (Pratt, 1990). This indicates that children can discriminate between 'good' and 'bad' questions, but their impulse to respond automatically takes precedence over comprehension monitoring in the moment. Ground rules should help reduce inappropriate responding by sanctioning a set of pre-ready responses children

can use instead, whilst also reminding them to monitor the confines of their knowledge and comprehension.

Finally, the immensurable range of abilities and stages of cognitive development children present with across childhood means any tool which gives an interviewer confidence in a child's narrative has value. As the cognitive skills required to recall episodic memories are rapidly evolving during child development, interviewers must adapt their approaches to help facilitate children's comprehension of their questions (Schneider, 2002). However, when interviewers question children, they cannot know for sure if the child has interpreted their question as intended. Research by Brown et al. (2013) found that even when interviewers were highly trained and closely followed the well-supported NICHD protocol (Lamb et al., 2018), 13% of the questions they asked were suggestive. Therefore, children need to have an accessible strategy to signal to interviewers when they have received a question they cannot answer, understand, or that is incorrect. When they use ground rules this way, interviewers can have more confidence in a child's narrative as it is assumed to be more accurate and reliable. As such, ground rules are believed to function as a way of priming children to recognise when there is something wrong with the interviewer's question and then how to respond appropriately.

So what do we know about ground rules from past research? In this study, we focused on three commonly used rules: I don't know ("If you don't know the answer to a question, you should say 'I don't know'": IDK), I don't understand ("If I [the interviewer] say something you do not understand, you should say 'I don't understand'": IDU) and correct me ("If I [the interviewer] say something incorrect, you should tell me what the correct thing is": CM). These rules form the focus of the following literature review.

Studies to date have examined ground rules in different contexts with respect to different outcomes. So far, there is no convergence across the literature to show how children's ability to use the rules change across development, how different methods of practice influence them and how they affect various measures of testimony quality. In addition to this, all the rules have not been studied to the same extent; the IDK rule has been studied the most extensively, and limited coverage exists of the IDU and CM rules (Brubacher et al., 2015).

Dickinson et al. (2015) have conducted one of the few studies using all three rules to understand developmental patterns in children's ability to apply them across childhood. They individually told 501 children (4-12 years old) a ground rule, then immediately after, asked them a question to see if they could correctly apply it. They found that children were relatively adept at using the IDK rule (phrased as "Don't guess") even from a young age. At age 4, 83.3% of

children could correctly respond to the IDK question, compared to only 31% for CM and 14.3% for the IDU question. This contrasts with Brown et al. (2019) who found that children across ages 4-12 ($N = 99$) were better at answering CM, rather than IDK questions immediately following ground rule instruction (IDU rule was not practised). Out of the three rules, the IDU rule appears to be the most difficult to understand, taking until ages 11-12 for children to reach 90% accuracy (compared to age 6 for IDK and age 8 for the CM rule) (Dickinson et al., 2015).

These developmental differences in the application of ground rules indicate that the cognitive abilities required to apply each ground rule develop at different rates. If this is true, then when a child displays competence in one rule, we cannot automatically confer competence in the other rules too. This was recently verified by Brown et al. (2019), who found that performance across the IDK, IDU and CM rules were not significantly related to each other. Furthermore, Brown et al. (2019) found that developmental differences strongly predicted children's ability to use ground rules in an interview context, and when answering practice questions following instruction. This adds further evidence that children may need to be supported in different ways to apply ground rules according to their developmental stage. Younger children are likely to need more instruction and practice than older children because of their relatively more immature social and cognitive skills. Unfortunately, many best practice protocols, including the NICHD protocol (Lamb et al., 2018), currently offer one-size-fits-all introductions for ground rules in interviews. Further research is therefore needed to confidently establish norms for how children's ability to understand and apply ground rules changes across childhood. This project will extend this research field by sampling from a wide age range (5-12) so that developmental trends may be accounted for in the analysis.

So does telling children of the ground rules help them in interviews? Ellis et al. (2003) examined whether instructing 4-5-year-olds ($N = 49$) to use the IDK and CM ground rules enabled them to resist suggestive questions when asked about an experienced event and a story they were read. The researchers staged an event that included three activities (two experienced and one story), then 1-2 days later, purposely introduced biased information to the children about two of the activities. Each child was interviewed using cued-recall and specific questions 8, 15 and 22 days later. At the start of each interview, half of the children were told the IDK and CM ground rules and the other half were told to try their best. No difference was found between the two groups on the number of false details children reported about the activities (Ellis et al., 2003). This pattern has also been observed in transcripts of police interviews conducted with 4-13 year-old suspected victims (sic) of sexual abuse (Earhart et al., 2014). In Earhart et al.'s

(2014) study, half the children were advised of the IDK rule at the start of the interview, and the other half were not. The results showed no difference between the two groups on the number of ‘I don’t know’ responses given in the interviews (Earhart et al., 2014). These studies suggest that merely telling children of the ground rules offers no added benefit over leading straight into the memory interview itself.

However, this has not been a consistent trend across the research literature. Mulder and Vrij (1996) found that telling children (4-5 and 8-10 years-old, $N = 109$) prior to the interview that ‘I don’t know’ was an acceptable response, led to a reduction of incorrect answers for difficult questions from 40% to 26%. Like Ellis et al. (2003), they interviewed the children about a staged event, but, their interview comprised of just one open-ended and three specific difficult questions. Additionally, it took place immediately following the event (Mulder & Vrij, 1996). These research conditions contrast widely from the contexts in which typical forensic interviews take place. In a forensic interview, children often have to recall event(s) which occurred weeks/months/years ago, and they are asked many questions, of varying types, in the interview. The importance of simulating these real-world conditions is underscored by the fact that question type (Waterman et al., 2000) and increased passage of time since the target event (Waterman & Blades, 2011) can influence children’s responses. Therefore, the lack of time delay and limited interview likely led to an over-estimation of the IDK rule’s effect in Mulder and Vrij’s (1996) study. Our research project incorporates these key factors by featuring a delay of 2-3 weeks between the event and interview and using the widely-adopted NICHD protocol to guide the interview structure (Lamb et al., 2018).

How, then, do children fare when they can practice the ground rules? When children are given one practice question following ground rule instruction, previous lab-based research has produced mixed results. Cordon et al. (2005) interviewed 3-6 year-olds ($N = 40$) about an experienced event and found that stating the IDK rule with an example of when to use it, and one practice question did not make a difference to how often they said ‘I don’t know’ in the interviews. More recently, Danby et al. (2015) found a different result with a larger sample ($N = 260$) of 5-9 year-olds. Each child participated in four scripted classroom activities and was interviewed about them 3-14 days later. At the start of the interview, they were advised of the IDK, IDU and CM ground rules and half got a practice question for each rule. The practice questions bore no semantic relation to the event, and if a child answered incorrectly, they were given feedback and up to two more attempts to answer the question. Children’s ability to apply the ground rules was assessed by three non-event-related difficult questions placed at the end of

the interview. In addition to rule use, Danby et al. (2015) also measured accuracy by assessing what proportion of 16 target memory items each child spontaneously recalled. Using this method, Danby et al. (2015) found that practice increased spontaneous use of IDK throughout the interview and in response to challenge questions, but there was no significant effects found in regard to children's accuracy or use of the IDU and CM rules (Danby et al., 2015). As Danby et al. (2015) used an older age group than Cordon et al. (2005), it could be that these differing results simply represent the relative developmental differences of the sample.

Although prior research has not asserted a clear benefit to offering just one practice question, this is precisely how forensic interview protocols often recommend that ground rules are introduced (Lamb et al., 2018; APSAC Taskforce, 2012). For example, in the NICHHD protocol (Lamb et al., 2018), children get one practice opportunity for the IDK and CM rules, whilst only instruction is offered for the IDU rule. As such, further research is needed to understand if children benefit from being introduced to ground rules in this way. Our research project extends this area of research by including a condition which broadly follows the NICHHD ground rules practice guidance (Lamb et al., 2018) (Business as Usual (BAU) condition). This enabled us to assess how the different practice methods compare to the status quo.

Past research has, however, found that practice does not always make perfect. When children are given more intensive practice for the IDK rule, they can become better at resisting difficult questions; however, this sometimes comes at the expense of reducing children's responses to questions that they can answer (Gee et al., 1999; Nesbitt & Markham, 1999; Saywitz & Moan-Hardie, 1994). For example, Saywitz and Moan-Hardie (1994) staged an event then two-weeks later, half the sample (7-years-old, $N = 102$) participated in a comprehensive practice regime. This regime involved listening to a story about a child who went along with leading questions, strategies for responding to difficult questions, practice at applying the techniques with feedback and a review session before the interview, which took place the next day. In that interview, the children in the training condition stated 'I don't know' more often and made less incorrect responses to difficult questions; but, they also produced fewer correct answers to non-misleading questions too (Saywitz & Moan-Hardie, 1994).

However, Gee et al. (1999) were able to rectify the over-generalisation of 'I don't know' responses with some simple additions to the training. In their initial study, half the sample of 9-13-year-old children ($N = 106$) practised the IDK rule with two wh-question types they did not know (e.g. What is my cat's name?), a wh- question they did know (What is your name?) and one option-posing question (Is my middle name Barbara or Jane?). Additionally, the interviewer

practised the CM rule with children by asking ‘Am I wearing scandals or gumboots?’ when they were obviously wearing shoes. Children who practised in this way were better at answering misleading questions than the Control group, but they also gave less appropriate responses to questions they could answer. In a second study (9-11-year-olds, $N = 51$), Gee et al. (1999) amended the practice method to provide greater reassurance there would be questions the children can answer, and that they do not have to say ‘I don’t know’ if they do know the answer. Additionally, they extended the ground rule practice questions to also include two non-misleading open and option-posing questions, in addition to the original four misleading questions. These adaptations decreased the errors to misleading questions successfully without also increasing error responses to non-misleading questions too (Gee et al., 1999).

Although we do not know which added component led to the children’s improved responses in Gee et al.’s (1999) second study, it did raise how ground rules practice may have broader effects on other important aspects of child testimony. We explore this topic further in our research by examining how different practice methods impact how accurate children are in a general NICHD interview (Lamb et al., 2018) in addition to how they respond to difficult questions. This allows us to capture a more holistic view of the impact of ground rule practice methods on interview quality. As this research question is exploratory, we do not have a specific prediction on the presence or direction of effect that the practice method may have on interview accuracy.

Across many of the existing studies conducted on ground rules, children are often assessed by how they respond to difficult ‘challenge’ questions inserted amongst other questions they know, understand or are correct. These challenge questions may mimic difficult questions interviewers ask in error (e.g., incorrectly paraphrasing a child’s earlier information), or because they do not know about the child’s experience (and so may ask a question that is beyond the boundaries of the child’s knowledge). How children answer these challenge questions allows us to directly assess whether they recognise when to use ground rules, and when they fail to do so. However, we currently do not know how children’s ability to use ground rules in response to challenge questions may impact other questions they can answer. On the one hand, children who are good at recognising difficult challenge questions may have improved general accuracy through better accuracy monitoring and reduced need for guessing, inferences or making assumptions (Brubacher et al., 2015). Alternatively, they might be less likely to answer questions that they should (as in the don’t know studies (Gee et al., 1999; Nesbitt & Markham, 1999; Saywitz & Moan-Hardie, 1994). This research project was designed to expand our

understanding by examining how children's use of ground rules in response to difficult challenge questions may also correspond to how accurate they are when reporting an experienced event in an NICHD protocol (Lamb et al., 2018) interview. As no previous research has assessed this relationship before, the author cannot predict the presence or direction of effects.

Past research on the IDU rule has found support for comprehension monitoring training as a method of boosting children's ability to detect incomprehensible questions (Peters & Nunez, 1999; Saywitz et al., 1999). For instance, Saywitz et al. (1999) staged an event for 6-8 year-olds (N = 180) then two weeks later, all children took part in one of three levels of comprehension training. The Control group watched video vignettes about the adverse effects of answering questions they do not understand and were told to indicate when they did not understand a question. The next level condition (Rephrasing Instructions group) received the same training as the Control group and instructions on requesting rephrasing of questions that they do not understand. The third condition (Comprehension Training group) got all the training mentioned above and practised detecting incomprehensible questions with feedback on performance (Saywitz et al., 1999). When each child was interviewed two days later, the Comprehension Training group correctly used the IDU ground rule 73% of the time, compared to 55% for the Rephrasing Instruction group and just 11% for the Control (Saywitz et al., 1999). This demonstrates that children potentially possess the capacity to monitor their comprehension beyond what their development profile would predict. In the 'real world', interviewers must weigh the benefit of taking more time to practice ground rules at the potential expense of a child's energy levels and motivation, not to mention logistical demands – all in a day's work. In acknowledgement of these concerns, each of the practice methods included in this study can be implemented within best-practice interview protocols already in use (Lamb et al., 2018; APSAC Taskforce, 2012).

If giving children multiple training sessions across numerous modes is not feasible for practitioners, then we need to find smarter and more efficient ways for children to practice ground rules. To this end, the author has looked outside the ground rules literature base to the Cognitive Learning and Education domains to understand how we can design a practice method that is both effective and practical. One learning theory which has received wide-support is 'structural alignment' (Braithwaite & Goldstone, 2015; Christie & Gentner, 2010; Day & Goldstone, 2012; Genter et al., 2007; Gentner, 1983). Structural alignment asserts that when we learn a set of concepts in one context and then successfully apply it in another context, this is

because we learned how the *relationships* between concepts are the same in each context, as opposed to how they are *literally* the same (Gentner, 1983). One way structural alignment can be enhanced is by increasing the match between the contexts where the relationship structure is learned and the context to which it should be applied to (Day & Goldstone, 2012; Gentner et al., 2007; Gelman et al., 2009). Braithwaite and Goldstone (2015) also found that this method works best when the person has limited prior knowledge of the concepts being learned. So what would this look like when trying to teach children the ground rules?

Firstly, we could ask practice questions like the types of questions they will encounter in the interview. Gee et al.'s (1999) second study did this with by including option-posing, open-ended, misleading and non-misleading questions in their ground rules practice. Perhaps not surprisingly, Gee et al. (1999) was also one of the few studies which found a positive effect in use of ground rules without a multi-session training regime. This provides encouraging evidence that more practice is not always what is needed, but instead, a practice which allows children to apply the ground rules in context can be just as, if not more, effective. Our research aims to extend upon Gee et al.'s (1999) work and past research in cognitive learning by administering different ground rules practice levels to children. These different practice levels vary by the match between the practice and interview contexts. One way the contexts are made more alike is by increasing the span of questions that the children can practice. As such, this research project includes a condition, Extended Practice (EP), in which children practice each ground rule with option-posing, wh- and open-ended type questions.

As children's answers can be influenced by question type (Waterman et al., 2000), practicing ground rules with the kinds of questions which commonly occur in interviews should enable children to recognise a more extensive range of questions they do not know, understand or are incorrect. Christie and Gentner's (2010) study supports this theory; they found that increasing the number of exemplars for comparison allows children to better detect what common relationships are shared amongst them. We therefore propose that increasing the match between practice and interview context will better facilitate children's use of ground rules when they encounter difficult questions.

One way which existing interview protocols try to give children the interview experience is via 'practice narratives'. Practice narratives are a 'warm-up' interview which asks the child about a recent non-target event (Roberts et al., 2011). Research has shown that including a practice narrative can positively influence the quality of child testimony by increasing the number of episodic details, amount and accuracy of information recalled in the latter substantive

phases of the interview (Brubacher et al., 2011; Price et al., 2013). As practice narratives are essentially ‘practice interviews’ it follows that embedding ground rule practice questions into the practice narrative could be another way ground rule practice can be more closely matched to the application context. As of yet, no research has explored this method; our condition with the closest match between practice and interview context therefore involves ground rules practice in the practice narrative in addition to the extended question types from the EP condition. The Extended Practice and Practice Narrative (EP+) method should theoretically draw children’s attention to the kind of questions they should use ground rules with and how they can arise within the interview structure. We predict that children who practice the rules this way will be the most competent at using the ground rules out of the four conditions.

This research project uses principles from Cognitive Learning and the practical constraints faced by interviewers to design two practice methods, along with a BAU and Control condition, that vary by match between the ground rules practice method and interview context. We note the lack of wide age ranges and the primary use of challenge questions to assess competence in past ground rules literature. These factors mean the broader effects of ground use on other aspects of child testimony has not been considered previously. Bringing this all together, we are interested in two questions: Firstly, does increasing the span of ground rules practice to more closely align with the interview context facilitate children’s use of the rules? And secondly, what are the broader effects of a) practice method, and b) ground rule use, on children’s general interview accuracy?

We draw from a larger research project that examined ground rule training approaches with children across development to answer these questions. Our study examines data from 5-12 year-olds interviewed about a staged event after a 2-3 week delay. Each child was assigned to one of four conditions which varied by the practice method used for ground rules. The first condition, the Control, did not receive any ground rule instruction or practice. Children in the BAU were introduced the ground rules as per the NICHD protocol (Lamb et al., 2018), and the New Zealand Specialist Child Witness Interview model (Lamb et al., 2018; Westera et al., 2016). Then in EP condition, children were asked a wider range of question types following ground rule instruction. The final condition, EP+, was built on the EP condition with practice questions also embedded in the practice narrative. Competence in using the ground rules was assessed by use of difficult questions designed to elicit the rules which were flexibly inserted in the memory interview. The details the children provided about the event were also coded for accuracy. We predicted that as similarity of the practice and interview contexts increased, the

children would become better at using the ground rules. No predictions were made about the effects of ground rule use and practice method in regard to accuracy as there was no current literature to guide us.

Method

This research project used a between-subjects design with children assigned to conditions that varied by the number and type of questions used in ground rules practice. The School of Psychology Human Ethics Committee approved this study under the delegated authority of Victoria University of Wellington's Human Ethics Committee (application #0000026259).

Participants

The Covid-19 Lockdown had a severe impact on the ability to stage events and interviews. From the initially planned sample of 288, just 93 children were included in this study. These 93 children assented and completed the Health and Safety event and interview before the Covid-19 Lockdown in March 2020. Data collection resumed in August; however, these children's data were unable to be processed and coded in time for inclusion in this study.

We recruited children between the ages of 5-12 from three Primary Schools in the Greater Wellington region, New Zealand. Each school was rated between decile 8-10, meaning they had less than 30% of children from lower socio-economic families. Children were randomly assigned to each condition and matched by school group, age and gender as far as possible; however, due to the data collection disruptions, children in the youngest and oldest age groups are underrepresented in the sample. See Table 1 for the final breakdown of age, gender and condition. Most children identified as New Zealand European (64.92%), 11.4% as Asian, 10.52% as Māori, 6.14% as European, 3.51% as Pacific Island and 3.51% identified with other ethnicities.

Procedure

Participating schools distributed consent forms to whānau/parents/caregivers of children in the desired age ranges. The children for whom consent was granted participated in two study sessions. Schools were thanked for their participation with a retail voucher calculated at \$1.30 per participating child (minimum of \$15.00).

Health and Safety Event

The children participated in a one-off Health and Safety event at their school, which was approximately 40 minutes long. They were allocated into groups of 6-8 by the Event Leader and took part in two of four activity stations focused on different aspects of health and safety.

Table 1*Number of Children in Sample by Age, Gender and Condition*

Age	Males	Females	Control	BAU	EP	EP+
5	3	6	3	3	1	2
6	6	9	4	5	5	1
7	6	5	4	3	3	1
8	10	8	7	4	4	3
9	5	13	4	7	5	2
10	2	8	4	2	2	2
11	3	4	4	2	0	1
12	4	1	3	1	0	1
Total	39	54	33	27	20	13

Research assistants, named Station Leaders, led the children throughout each activity. Children attended either the Hazards and Heartbeat stations (H&H) or the Treatment of Cuts and Temperature stations (T&T). Important information not captured in the script such as who the children's partners were, deviations from the script and any unusual child behaviour were recorded and later used to verify interview details.

While the children were at the second station, a (previously unseen) research assistant entered the room and asked each Station Leader where they could find the Event Leader. At this point, the Station Leader would stop and pay attention to the unfolding situation, which naturally cued the children to also observe the interruption. The research assistant then proceeded to have a heated discussion with the Event Leader about how they were unaware of the current event taking place and needed supplies to deliver another Health and Safety event elsewhere. The research assistant and the Event Leader resolved the disagreement and agreed the interrupter could use the spare equipment instead. The interrupting research assistant then collected the 'spare equipment' from each Station Leader and left the room, after which, the station leaders resumed their respective activities.

Each child received a prize (pencil) to thank them for their participation at the end of the event. Each station is summarised below, and the full event script is in Appendix A.

Treatment of Cuts Station. Children watched a cartoon video about a boy named Arthur who cut his knee while trying to get a tyre for his bike at the local dump. Following the video, the Station Leader asked questions about what happened to Arthur and how they could

have helped him. After the discussion, the leader drew a pretend cut on each child's left index finger then presented a PowerPoint demonstrating step-by-step to the children how to care for a cut. Together with the leader, the children practised each step on their pretend cuts: applying pressure, raising it above their heart, cleaning it with an antiseptic wipe and putting a plaster on it. The leader then took a picture of the children holding up their fingers with the plasters on, which concluded the station.

Temperature Station. The Station Leader told how temperature could indicate health and explained that measuring temperature in different places on their body will give different readings. The children were then placed in pairs and practised taking each other's temperature on their foreheads with the back of their hands. Then they used a stick thermometer to measure their partner's temperature under their left arm and left knee. The thermometers were sterilised with an antiseptic wipe whenever the children swapped turns. Children were then instructed that the best place to take the temperature was in the ear, and they took their partner's temperature in their left ear using a digital thermometer. They recorded their ear temperatures on a chart.

Hazards Station. At this station, the Station Leader showed children a series of coloured posters that depicted various hazards around the home. First, children were shown an example poster then at least five more. Additional posters were displayed if there was spare time. The Station Leader asked the children to identify critical hazards and say how they could make it safer. This discussion was recorded using a voice recorder device and later transcribed by a lab research assistant.

Heartbeat Station. At this station, the research assistant gave the children an explanation about why we listen to the heart and how it indicates fitness and health. The leader then showed them the stethoscopes covered in a range of animal covers (crocodile, bear, tiger). In pairs, the children listened to each other's chest and stomach with the stethoscopes. Between turns, each child sterilised the stethoscope earpieces with an antiseptic wipe. Then they felt their partner's pulse on their wrist and their ankle. At the end, the leader took a group photo of the children holding their stethoscopes.

Interview

Between 14-26 days following the Health and Safety event, each child was individually interviewed at their school by one of two female research assistants who were not present at the event. Interviews took between 20 to 55 minutes, with the average interview lasting 34.21 minutes ($SD = 7.88$). Before the interview started, children gave written (ages 9-12) or verbal consent (ages 3-8) and advised they could leave at any time. Interviewers were trained to follow

the NICHD protocol (Lamb et al., 2018), which relied predominantly on open-ended questions and children's utterances to prompt further elaboration. More specific and option-posing questions were used as needed to elicit important information and clarify ambiguous responses. When these were used interviewers then returned to an open prompt for their next question.

Children in each group except for the Control had three distinct components to the interview which were completed in the following order 1) ground rule instructions, 2) practice narrative, and 3) memory interview of health and safety event. The type of ground rules instruction and practice children received varied across four conditions which they were randomly assigned to. Each condition (described below) varied by degree of match between ground rules practice method and interview context.

Control. In this condition, no ground rules instruction or practice took place; instead, children began the interview with the practice narrative.

Business as Usual (BAU). Children were initially told they should only tell the truth and asked two option-posing questions which they had to answer as being true or not (e.g. *"If I say that my shoes are red, is that true or not true?"*). Then they were told of each ground rule (IDU, IDK, CM) and asked two wh- type questions to practice each rule. The interviewer praised correct answers, and incorrect answers were given feedback and reminded of the rule. Children got two opportunities to practice each rule, see Figure 1 for an illustration of the process. Examples of the practice questions can be found in Table 2 and the full list is in Appendix C.

Figure 1

Format of Practice for Each Ground Rule Per Condition

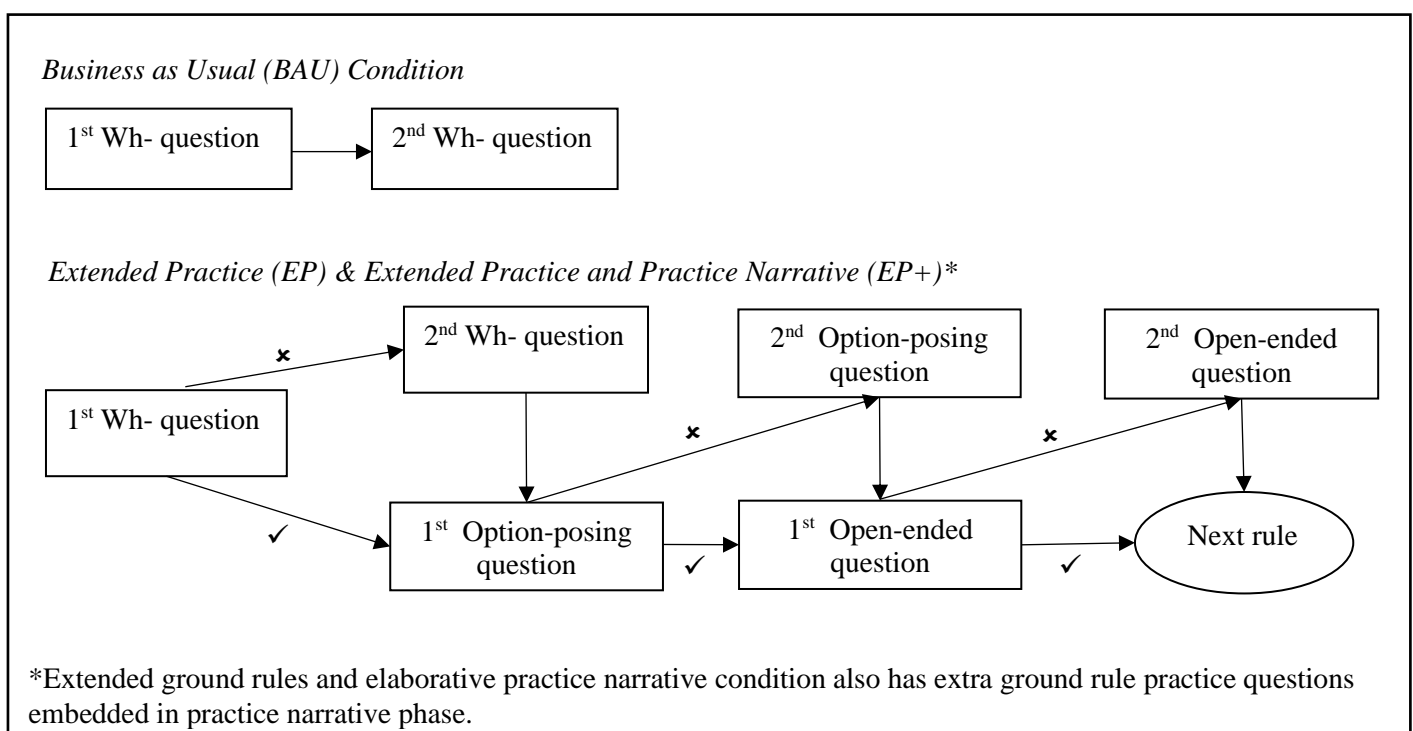


Table 2*Examples of Ground Rule Practice Questions*

Rule	Question type		
	Open-ended	Directive (Wh-)	Option-posing
‘I don’t know’	“Tell me all about my last birthday party.”	“What is my dog’s name?”	“Did I go to the movies last night?”
‘I don’t understand’	“Tell me all about your curmudgeon teacher.”	“What are you most adroit at?”	“Is my shirt gridelin?”
‘Correct me’	“Tell me about the Prime Minister’s visit to your school/preschool this morning.”	“Where in Australia do you live?”	“Do you live in Christchurch or Dunedin?”

After delivery of the rules, the interviewer gave a final reminder of each of the rules then proceeded to the practice narrative.

Extended Practice (EP). This group followed the same procedure as BAU but with extended question types when practicing the ground rules. Children got to practise telling the truth with an extra option-posing question and they practised each ground rule with option-posing and open-ended questions too. As with the BAU condition, if they got the question right, they were praised and went on to the next question. Incorrect answers were given feedback and then asked another same-type question. As such, children in this condition had up to two opportunities to answer an option-posing, wh- and open-ended practice question for each ground rule. In total, children were asked 9-18 practice questions depending on how many initial questions they got wrong – see Figure 1 to illustrate.

Extended Practice and Practice Narrative (EP+). This condition built upon the EP condition with one option-posing, wh- and open-ended practice question for each ground rule interspersed in the practice narrative, see Table 3 for examples. This provided extra opportunities for the child to practice the ground rules within a more authentic interview context. Interviewers flexibly inserted these questions in the practice narrative to adhere to the direction of the child’s narrative. Three versions of the practice narrative practice questions were counterbalanced across children in this condition – see Appendix C for each version. If the child responded to practice questions in the practice narrative with the appropriate ground rule, they received praise. If they asked the meaning of a word instead of applying IDU rule, the

interviewer acknowledged there was a tricky word in the question and then moved on. If the child responded to a practice question with the wrong ground rule, they were asked to explain why they said that. Children were also questioned if they attempted to answer an IDK, IDU or CM question without using a ground rule, (e.g. for IDK “*Do you really know [what happened when...]?*”), and then reminded of the appropriate ground rule. If a child guessed the meaning of a tricky word used in an IDU practice question or did know the answer to an IDK practice question, the interviewer acknowledged their correct answer then reminded them of the ground rule. At the end of the practice narrative, each child was reminded of the ground rules a final time before beginning the memory interview.

Table 3

Examples of Practice Questions Used in the Practice Narrative

Rule	Question type		
	Open-ended	Directive (Wh-)	Option-Posing
‘I don’t know’	“Who is your best friend? [wait for an answer] Tell me everything that happened when they got up this morning?”	“How many people were walking on your street before you woke up this morning?”	“Did your teacher eat yogurt for breakfast this morning?”
‘I don’t understand’	“Tell me about feeling acrasial this morning.”	“How did you make your bag pendant?”	“Is [just disclosed detail] a hemerine part of your morning?”
‘Correct me’	“Tell me everything that happened when the fire engine came to school this morning?”	“What colour was the hot air balloon that was in the sky on your way to school?”	“Did you go back home to get the school bag you forgot this morning?”

In the practice narrative of each condition, the interviewer asked the child to tell them everything they could remember from when they woke up until they arrived at school that morning. Children answered a range of questions in the practice narrative including general invitation questions (e.g. *Describe some other things that happened this morning*”), cued

invitations (e.g. “*So you were telling me about [detail child mentioned]. Tell me all about when that happened*”), wh-questions with pairing (e.g. “*What was the very next thing to happen?*”)

Table 4

Examples of Challenge Questions Used in Memory Interview

Rule	Question type		
	Open-ended	Directive (Wh-)	Option-Posing
‘I don’t know’	“Tell me everything that happened with the leaders before they came and got you.”	“The woman who came in, what colour was her bicycle?”	“You know the girl in the slideshow who cut her finger, was her name Sarah?”
‘I don’t understand’	“Tell me about measuring febrility?”	“Who auscultated your pulse?”	“Did Arthur hurt his patella?”
‘Correct me’	“Earlier you said you did different activities, tell me about what you did at the [wrong station] hazards station.”	“Where was the leader wearing her Bear stethoscope?”	“Did you use a red pen to draw a cut on your knee?”

then after answer, “*Tell me some more about [detail mentioned in previous answer]*”), option-posing questions with follow-up pairing (e.g. “*Did that happen before you left for [school/preschool]?*”) then after answer, “*Tell me some more about [same probed detail]*”), and summary invitations (e.g. [short summary of what child said in their words] then “*Tell me any other things you can remember about what you did this morning*”). After the practice narrative children were told,

“Now that I know you a little bit better let me tell you why I’ve come to talk to you today. I heard that a couple of weeks ago, some people came to your school, and you did some health and safety activities. Now, I wasn’t there, so I don’t know what happened, but I’d like to know all about it.”

Each memory interview started with general invitation questions to initiate free recall (e.g. “*Tell me everything that happened from the beginning to the end*”) followed by targeted recall follow-up questions (e.g. “*Earlier you mentioned [station already mentioned by the child using their own words]. Tell me everything you can remember about when that happened*”).

Challenge questions designed to elicit the ground rules were flexibly interspersed throughout the memory interview to ensure natural flow with the child's narrative. Three sets of challenge questions were used to counterbalance questions across the interviews, see Table 4 for examples and a full list is in Appendix B. After the interview, each child was thanked and asked not to talk it with other children at school.

Data Processing

Each interview was transcribed verbatim by trained research assistants from digital video recordings. The author transcribed 27 of these. Overall, 114 transcripts were coded, which included interviews from another condition not used in this study but formed part of the broader research programme. Two research assistants coded the responses to ground rule challenge question whilst the author and another research assistant coded units of information (UOI) and accuracy. Coding partners resolved disagreements with each other through discussion.

Responses to Ground Rule Challenge Questions

Responses were coded and scored by whether children used the correct rule, resisted answering the question or gave an inappropriate response, see Table 5 for sample responses. An initial 15 transcripts were used as training examples to test the coding scheme, then 20 transcripts were independently coded to assess inter-rater reliability using Cohen's κ . There was almost perfect agreement (Landis & Koch, 1977) between the two coders across children's responses, $\kappa = .925$ (95% CI, .909 to .941, $p < .001$). The remainder of the transcripts were then independently coded (and the training transcripts re-coded), with unusual responses coded by discussion. Eleven transcripts were reserved and periodically coded by both coders to reassess inter-rater reliability. Cohen's κ remained consistently high between the two coders throughout (Second check $\kappa = .933$, 95% CI, .917 to .949, $p < .001$; Third check $\kappa = .932$, 95% CI, .918 to .946). During coding, it was discovered that some children were not asked all the ground rule challenge questions due to interviewer error. Therefore, a ratio was calculated for each child of the score they received for that rule, divided by the total potential score that they could have received if they answered every challenge question for that rule correctly. These ratios functioned as the key dependent variables for challenge question performance.

Units of Information (UOI) and Accuracy

Details offered by children that were related to the Health and Safety event were as coded as UOI. UOI included people that were there, actions done by individuals, objects, locations, order/timing of events and general descriptions of event aspects. For example, a child who said "I measured my partner's temperature with two measuring sticks" would receive five

UOI: “I [measured] my [partner’s] [temperature] with [two] [measuring sticks]”. Initial responses to ground rule challenge questions, repeated details or responses not related to the event were not coded. UOI were then coded as being accurate, inaccurate or unverifiable. Unverifiable UOI included subjective evaluations made by the child and events not captured in the script which could not be substantiated as accurate or inaccurate. The number of accurate UOI said by the child was divided by the total number of accurate and inaccurate UOI they said (unverifiable UOI were excluded) to produce an accuracy score. This ratio was used as the dependent variable for overall interview accuracy.

Table 5

Response Descriptions to Ground Rule Challenge Questions with Associated Scores

Description	Score	Response Examples
		I: “The woman who came in, did she arrive by bicycle?” C: “Uhh, I don’t know”
Correct response	2	I: “Where was the man who came in going?” C: “It was a woman.”
		I: “Did you put the thermometer next to your tympanum?” C: “I don’t understand”
Resisted response or use of the wrong rule	1	I: “When did the bellicose woman leave the classroom?” C: “Mm, I don’t know”
Incorrect response	0	I: “Tell me about measuring febrility.” C: “It was really fun we did it in pairs.”
Abstained response	0	I: “Tell me about measuring arrhythmia.” C: “After the event, we got pencils as a prize.”

The coding process for UOI and accuracy took a total of four months to complete. During this time, a coding system was developed, 18 transcripts were used to finalise the coding scheme, and 38 transcripts were then coded to assess inter-rater reliability. There were very few instances of ‘inaccurate’ codes, so these were amalgamated with the unverifiable codes to calculate Cohen’s κ for accuracy. There was substantial agreement (Landis & Koch, 1977) between the coders for UOI, $\kappa = .734$ (95% CI, .718 to .75, $p < .001$), and almost perfect agreement (Landis & Koch, 1977) for how accurate they rated the units, $\kappa = .819$ (95% CI, .801 to .837, $p < .001$). Though Cohen’s κ agreement for UOI was less than that for accuracy, they had similarly high inter-rater percent agreement across the transcripts used for calculating reliability, 87.41% for UOI and 92.66% for accuracy.

After reliability was established between the two coders, each coder then coded another 39 transcripts and re-coded 16 of these to ascertain reliability with the self. Cohen’s κ was used to determine agreement across transcripts for each coder. The author had almost perfect agreement (Landis & Koch, 1977) with themselves for UOI, $\kappa = .879$ (95% CI, .863 to .895, $p < .001$) and for accuracy, $\kappa = .886$ (95% CI, .862 to .909, $p < .001$). The other coder similarly had almost perfect agreement with themselves (Landis & Koch, 1977) for UOI, $\kappa = .866$ (95% CI, .84 to .891, $p < .001$) and for accuracy, $\kappa = .91$ (95% CI, .939 to .881, $p < .001$).

Results

IBM SPSS Statistics and R statistical programmes were used for the data analysis. As the ability to use one ground rule does not necessarily indicate competence in using other rules (Brown et al., 2019; Dickinson et al., 2015), the children’s challenge question performance was first broken down by rule type.

Bivariate correlations were conducted to ascertain if the analysis should utilise children’s overall challenge question performance or only those measuring the same rule type. Children’s scores for the CM and IDU rules were weakly positively correlated but only to marginal significance (see Table 6). The IDK rule was not significantly associated with any rule. As the correlation between the IDU and CM rules was only marginal (and weak in strength), children’s performance to the challenge questions was analysed by rule type instead of overall responding.

Descriptive statistics (see Table 7) were conducted, and the data for children’s challenge questions performance across each rule appeared normally distributed with acceptable ranges of skew and kurtosis (Field, 2018). However, the accuracy data was negatively skewed with high kurtosis, in addition to having four outliers. As the normality assumptions were violated for

accuracy, and because proportional data was used, the accuracy ratio variable was arcsine transformed.

Table 6

Correlations of Children's Performance on Challenge Questions

Rule type	1	2
1. 'I don't know'	-	
2. 'I don't understand'	.11	-
	$p = .294$	
3. 'Correct me'	.06	.18
	$p = .582$	$p = .081$

Following transformation, skewness (-0.53 , $SE = 0.23$) and kurtosis (1.39 , $SE = 0.45$) now fell within an acceptable range (Field, 2018). The outliers were kept in the data due to the small sample size and corrected distribution following the arcsine transformation. The arcsine transformed ratio is used in the rest of the analysis, except for descriptive statistics.

Across the different rule types, children on average used the rules to a similar extent, with only a 5% difference separating the most difficult (CM) from the easiest (IDU) rule. On average, children were highly accurate in their general interview, which explains the negatively skewed distribution. Interestingly, the least accurate child was better than chance, and the most accurate child reported no inaccurate details at all.

Table 7

Descriptive Statistics of Challenge Question Performance and Accuracy

Rule	Mean	SD	Minimum	Maximum	Skewness	Kurtosis
'I don't know'	.6	.3	0	1	-0.15	-1.05
'I don't understand'	.63	.31	0	1	-0.38	-.927
'Correct me'	.58	.28	0	1	-0.17	-.84
Accuracy	.9	.08	.54	1	-2.05	5.49

Note. These descriptive statistics are reported from the raw data.

Preliminary Analysis

Independent t-tests were conducted to understand if there was any effect of gender on challenge question performance or accuracy. No significant effects were found across any of the rules (IDK = $t(91) = .29, p = .773$, IDU = $t(91) = 1.36, p = .177$, CM = $t(91) = 1.28, p = .203$) or accuracy ($t(91) = -.27, p = .785$). Similarly, no effect of interviewer was found on challenge question performance (IDK = $t(91) = -.24, p = .813$, IDU = $t(91) = -.61, p = .545$, CM = $t(91) = -.69, p = .492$) or accuracy ($t(91) = -.52, p = .605$).

Bivariate correlations were conducted to examine if duration of interview or delay between event and interview significantly influenced children's performance on the dependent measures. No significant relationships were revealed between duration of interview and challenge question performance (IDK = $r = .04, p = .699$, IDU = $r = .85, p = .418$, CM = $r = .11, p = .311$) or accuracy ($r = -.16, p = .126$). Additionally delay (days) between initial health and safety event and interview was not related to IDK ($r = -.01, p = .921$) or IDU ($r = .12, p = .862$) challenge question performance but it was significantly weakly correlated with responses to CM challenge questions ($r = -.25, p = .015$). Consequently, delay was included as a covariate in subsequent analysis for CM challenge question performance. No significant association was found between delay and accuracy, $r = -.13, p = .23$.

A one-way Analysis of Variance (ANOVA) was conducted to ascertain if the practice conditions were equivalent concerning age, and no significant effect was found ($F(3, 89) = 0.8, p = .497$). A chi-square test was also used to assess if the practice conditions significantly differed in relation to gender, and no significant difference was found, $\chi^2(3) = .94, p = .815$.

Main Analysis

Does Increasing the Match Between Ground Rules Practice and Interview Context Facilitate Children's Use of the Rules?

'I Don't Know' Challenge Question Performance. When the means were examined (see Table 8) from the different practice conditions, it appeared that on average children were best at recognising difficult questions they did not know when the practice and interview conditions were most closely matched (EP+). However, children in the Control condition surprisingly had the second-highest level of average performance. This could imply that the addition of practice questions in the practice narrative offered marginal benefit.

At each level of the independent variable, the data was inspected for the assumptions of normality, presence of outliers, homogeneity of regression slopes and homogeneity of variance.

Table 8*Descriptive Statistics of Challenge Question Performance by Rule and Condition*

Condition	Rule					
	‘I don’t know’		‘I don’t understand’		‘Correct me’	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Control	.61	.3	.43	.3	.51	.28
BAU	.56	.34	.71	.26	.58	.29
EP	.59	.25	.71	.29	.66	.28
EP+	.68	.28	.87	.14	.63	.23

No violations were found. An Analysis of Covariance (ANCOVA) was conducted to determine if there were statistically significant differences across the match between practice and interview context on the children’s performance on IDK challenge questions while controlling for age (years). The covariate, age, was not significantly related to children’s performance on IDK challenge questions ($p = .43$). Additionally, no significant effect was found for the match between practice and interview context on the way children responded to IDK challenge questions when controlling for age, $F(3, 88) = 0.43$, $p = .729$.

‘I Don’t Understand’ Challenge Question Performance. The descriptive statistics show that on average, children seemed to improve at answering questions they did not understand when practice context became more progressively matched to the interview context; but there was no difference in performance across the BAU and EP conditions. There was a 44% difference between the worst-performing (Control) and the best performing group (EP+), which suggests that substantial gains could be obtained by practising the IDU ground rule with children.

Data at each level of the independent variable was found to meet the assumption of no outliers but violated the assumptions of normality, homogeneity of regression slopes and homogeneity of variance. A log transformation of the data did not rectify the assumption violations, so a robust linear regression (Schmidt & Finan, 2018) was conducted with practice condition and age used as the explanatory variables and score for IDU challenge questions the outcome variable. The residuals were normally distributed, which indicated the appropriateness of the model for this data.

The regression results indicated that the predictors significantly explained 29.23% of the variance ($R^2_{\text{adjusted}} = .26$). The covariate, age, marginally significantly predicted children’s

performance on the IDU challenge questions, $\beta = .16, p = .081$ when the practice condition was controlled for. Therefore, for every one-year increase in age, children's performance would improve by 16 percentage points. Furthermore, a significant difference was found between children's performance on IDU challenge questions across the different practice spans, $F(4,88) = 9.08, p = <.001$. The estimated marginal means for each group were extracted from the linear model and a significant difference was found between the Control group (no practice) ($M = .42, SE = .04$) and the BAU ($M = .71, SE = .05, p = <.001$), EP ($M = .72, SE = .06, p = .001$) and EP+ conditions ($M = .87, SE = .07, p = <.001$). None of the other practice conditions were significantly different from each other for the IDU rule.

‘Correct Me’ Challenge Question Performance. From the descriptive statistics, we can see that children's performance on CM challenge questions was the lowest amongst the Control group and the highest in the EP condition. This suggests children benefitted from the extra range of question types offered in the EP condition, but not from ground rules practice in the practice narrative.

The assumptions of normality, no outliers, homogeneity of regression slopes and homogeneity of variance were met for children's performance on the CM challenge questions. Both covariates, age and delay between the event and interview, were significantly related to children's performance on the CM challenge questions, age $F(3,87) = 10.96, p = .001$, delay $F(3,87) = 11.96, p = .001$. However, no significant differences were found between children's performance across the different practice spans when age and delay between event and interview were accounted for $F(3, 87) = 1.61, p = .194$.

Summary

The statistical analysis suggests that when children have to practice ground rules, increasing the match between the practice and interview context does significantly affect children's use of ground rules when being asked difficult questions. Furthermore, it indicates that introducing ground rules (with any level of practice) offers no benefit over simply starting the interview with a practice narrative, when it comes to helping children answer questions they do not know or that are incorrect. However, when children are asked questions with words they do not understand, any form of practice can help them be better at answering these types of questions, compared to children who do not get any introduction to ground rules in the interview.

Does the Match Between Ground Rules Practice and Interview Context Affect How Accurate Children Are When Reporting a Recent Event?

Concerning accuracy, children were similarly accurate on average across conditions (see Table 9), with only a 3% difference separating the least accurate (EP) from the most accurate group (EP+). Interestingly, the children in the EP+ condition also had the smallest range with all children being between 85-97% accurate. This suggests that although children are on average similarly accurate irrespective of training type, they may be more consistently accurate when they can practice the ground rules in a way that closely matches the context they should apply them.

The data was inspected across each of the practice conditions, and the assumptions were met for normality, presence of outliers, homogeneity of regression slopes and homogeneity of variance. An ANCOVA was performed to determine if the practice condition significantly affected children's overall accuracy when controlling for age. No significant effect was found, $F(3,88) = 0.11, p = .954$. This indicates that how children learn and practice the grounds (or not) has no significant flow-on effect to how accurately children describe a recent memory event.

Table 9

Descriptive Statistics of Accuracy Ratio by Condition and Overall

Condition	<i>M</i>	<i>SD</i>	Minimum	Maximum
Control	.9	.08	.61	.97
BAU	.9	.06	.72	1
EP	.88	.13	.54	1
EP+	.91	.04	.85	.97

Is Ground Rule Use Associated with the Accuracy of Details Children Report About a Recent Event?

Partial correlations were conducted to examine if children's responses to challenge questions across the different rule types were correlated with overall interview accuracy when age was controlled (Table 10). No significant correlations were found between any of the measures. This implies that how good (or not) a child is at applying ground rules in response to difficult questions is not significantly associated with how accurate they are when reporting about a recent memory event.

As the partial correlations did not yield any significant relationships between ground rule use and accuracy, no further analysis was conducted.

Table 10*Correlations of Accuracy and Challenge Question Performance Whilst Controlling for Age*

Rule	Interview Accuracy
'I don't know'	.12 $p = .266$
'I don't understand'	.09 $p = .41$
'Correct me'	.12 $p = .257$
All rules	.17 $p = .102$

Discussion

This research project examined the match between practice and interview contexts and its influence on children's use of the IDK, IDU and CM rules when asked difficult questions. This project also explored how the practice methods and use of the rules affected children's accuracy when reporting about a recently experienced event. The research did this by staging an event for children at their school and interviewing them about it 2-3 weeks later. Before the memory interview, children practised the ground rules in one of the three ways varied by the match of practice with the interview context or were not introduced to the rules at all (Control condition). Children's responses to ground rule challenge questions in the interview and the accuracy of details they recalled about the event were coded and used as our key dependent variables. Before the results are discussed in-depth, the author acknowledges the external influence the Covid-19 pandemic and associated Lockdowns had on the final number of participants included in this study. These results should therefore be interpreted with some caution.

Does Increasing the Match Between Ground Rules Practice and Interview Context Facilitate Children's Use of the Rules?

We predicted that as children increasingly practised the ground rules in a similar context to which they are expected to apply them, we would see increased use of the rules in response to difficult questions. Additionally, we predicted that the condition with the highest degree of match between practice and interview context (EP+) would also have the highest level of rule use out of all the conditions. We found no significant differences in how children responded to difficult questions on any of the rule types across any of the conditions which included some level of ground rules practice (BAU, EP, EP+). Therefore, neither of our predictions were

supported. The only significant effect that emerged was that for the IDU rule; children in the Control group were significantly worse at answering incomprehensible questions than in the other practice conditions (43% and 71% respectively).

At face value, these results suggest that the match between the practice and interview context is not a significant factor in helping children to recognise and respond to difficult questions, regardless of rule type. Furthermore, the findings indicated that even introducing the ground rules with any level of practice does not offer any significant benefit when children are asked questions they do not know or that are incorrect. Conversely for IDU, they suggest that any form of practice is enough to help children apply the IDU ground rule when they encounter questions they do not understand.

These results are a concern for practitioners following many forensic interview protocols. For example, in the NICHD protocol (Lamb et al., 2018) only the IDK and CM rules are practised, and the IDU rule is delivered as an instruction. The evidence collected in our research strongly questions this method (for children 5-12 years old); however, more research is still needed to understand how children's baseline ability to apply ground rules changes across childhood.

What possible explanations can be made about the pattern of results we found? One probable viewpoint is that children already bring a mastery of the IDK and CM rules to the interview context. Dickinson et al. (2015) and Brown et al. (2019) found that young children were relatively competent at applying the IDK and CM rules, becoming at least 90% accurate at answering IDK practice questions from age six and CM from age eight. In Dickinson et al.'s (2015) it took children until ages 11-12 to reach the same accuracy level for IDU questions. Additionally, we know children across the ages 5-11 commonly use the phrase 'I don't know', for a variety of reasons (Vogl, 2012), e.g. they are feeling indifferent, or it may be too taxing for them to have to generate an answer. As such, 'I don't know' is a readily accessible response that children already commonly use when responding to difficult questions. These trends were reflected in our data, with the Control group getting an average of 61% of IDK and 51% of CM challenge questions correct, but only 43% of IDU challenge questions correct. Therefore, it could be that we observed ceiling effects regarding children's performance on the IDK and CM rules. These ceiling effects would have prevented any meaningful variation from emerging between the practice conditions for these two rules.

Additionally, it appears that merely introducing the IDU ground rule with at least one practice question is enough to make incomprehensible questions more salient to children. Our

results, along with the trends seen in previous literature, call into question if children need to practice the IDK and CM rules at all. It might be more efficient for interviewers to only practice the IDU rule; conserving children's energy and motivation for the rest of the interview.

These relative differences in difficulty across the ground rules are likely to be connected to the cognitive processes involved in applying ground rules in response to difficult questions. When we examine the three rules from a cognitive standpoint, we note that the IDK and CM rules are similar in that for children to apply them, they must monitor the knowledge of their own experience. However, to use the IDU rule, children must monitor a different aspect of memory, their lexicon. To do this, a child has to consciously interrupt an automatic process called 'bootstrapping' which occurs when they encounter unfamiliar words (Höhle, 2009). During bootstrapping, children infer meanings of words they do not know by how it is placed within a sentence, and it is one of the critical mechanisms involved in language acquisition (Höhle, 2009).

Although children commonly encounter words in their everyday lives that they cannot understand, it would be uncommon for them to explicitly indicate when this is the case. As such, they are unlikely to have much practice interrupting the automatic process of bootstrapping and in turn, find applying the IDU rule a demanding task. In contrast, children have much more practice at recalling events, allowing them to more easily discern the boundaries of their memory. Fortunately, our study provides some evidence that instructing children to be aware of questions they do not understand and giving them a question to practice with, is enough to boost their ability to overcome the bootstrapping process.

Another potential reason why no significant effect was found across the different practice conditions could be due to how children experienced the event and interview. In our study, the children were actively engaged throughout the activities and likely experienced it as novel with a relatively positive emotional valence. Although children were instructed to not talk about the event with their peers, it's highly likely that many of them still rehearsed the event with friends and family members, e.g. when asked by their parents about their day at school. Additionally, the interview took place only 2-3 weeks later, in a familiar environment connected to where they experienced the event (their school) and was with a highly-trained interviewer. All these factors contribute to optimum encoding and later recall ability (Lamb et al., 2018). This was evidenced by the fact that children were highly accurate in recalling details from the event, averaging between 88-91% accuracy across the different conditions. This high accuracy in recall suggests that the Health and Safety event was well-remembered; therefore, this could have led to ceiling

effects in the IDK and CM challenge questions because children could monitor the confines of their memories more easily.

To overcome these potential ceiling effects, future research should, as far as possible, simulate the conditions children experience in forensic interviews for more authentic responses. When people recall events of high emotional arousal and negatively emotionally valenced, a distinct neural pathway is activated compared to when they recall memories that are neutral in arousal and emotional valence (Kensinger & Corkin, 2004). If different cognitive processes are used to recall distressing memories, we may see children use ground rules differently in these conditions. As it is difficult to ethically create these conditions in lab-based environments, future research could benefit through the use of more field-based data to understand the influence of a child's cognitive state on how they use ground rules in interviews.

When the children were confronted with the challenge questions designed to elicit the IDU rule, it is possible the language of these questions may have made them more salient compared to the other rule type challenge questions. This is because the only time the children encountered words they potentially could not understand was when they were asked IDU challenge questions. This may have differentiated the IDU challenge questions from all the other questions, which used simple language. As such, alerting children to incomprehensible questions in the ground rules instruction and practice could have made 'I don't understand' a more accessible response when they encountered a question they did not understand. This may also explain the lack of difference between the groups which received some form of practice as they all received the same initial instruction of the ground rules, albeit with varying levels of practice. The author recommends that future research be should aware of this potential limitation and try to find ways to counter-balance it where possible.

Finally, it could be that increasing the match between practice and interview context does not provide any noticeable benefit to helping children apply ground rules. But why might structural alignment not work for ground rules? In Braithwaite and Goldstone's (2015) study they gave half their participants a pre-instruction on how to search for specific structural features, and the other half received no extra instruction. They found that the group who received the pre-instructions did not benefit from increasing the similarity of the test items whereas the group who did not receive pre-instruction was able to more easily recognise common structural features in logic problems when the test items looked superficially similar. Although children have limited exposure to using ground rules in forensic interviews, they do have significant experience of encountering questions they do not know, understand or that are

incorrect in their everyday life. This prior knowledge may have negated any effect created by increasing the match between the ground rules practice and interview contexts. To assess the viability of this explanation, future research could assess how children's proficiency at using ground rules at differs across different levels of match between practice and interview context when compared to other practice techniques.

Does the Match Between Ground Rules Practice and Interview Context Affect How Accurate Children Are When Reporting a Recent Event?

No significant differences were found between the different practice conditions and how accurately children were able to recall the event (outside of the challenge questions). Furthermore, it is interesting that no significant differences were found between the Control group, who received no ground rule instruction, and the other practice conditions. This suggests that even introducing the ground rules has no flow-on effect to accuracy with any level of practice. The absence of adverse effects on accuracy may reassure practitioners. However, it may also indicate there is no positive impact when including ground rules instruction and practice (compared to no ground rules instruction and practice) on how likely children report false details. Why would instructing children of the ground rules not also alert them to monitor their narratives' accuracy?

Firstly, it could be that ground rules only help when interviewers ask tricky, as opposed to well-constructed, questions. Outside of the times that children were asked the challenge questions, interviewers strictly followed the NICHD protocol which relies primarily on open and cued-recall questions to elicit details about the event (Lamb et al., 2018). We know that children tend to be reliably accurate when they are interviewed this way (Lamb et al., 2018), and this was reinforced by the fact that children were highly accurate across each of the conditions. Due to the strict adherence to the NICHD protocol, it's possible that we observed ceiling effects in the accuracy measure. These conditions are unlikely to replicate in actual forensic interviews where divergence from best practice is commonly observed (Lamb et al., 2018). Perhaps further investigations might see an association amongst children who are good at using ground rules in response to close-ended questions, or questions that ask for more elaboration about a detail, or those which occur later in the interview when a child's memory is more likely to be taxed. It is recommended that future research consider accuracy with different categories of question type to understand if, under more authentic conditions, an interaction between accuracy and ground rules may emerge.

Another explanation for the results obtained could be that match between ground rules practice and interview context increased accuracy monitoring, but only on what children did *not* say. For example, children exposed to ground rule instructions and practice could have been editing their accounts before verbalising them, and only sharing information they felt confident about. Past research has indicated that children frequently do not tell us all they know in interviews, even when asked. This is evidenced when researchers ask children recognition questions at the end of an interview, and children then report on new details, with elaboration, that were not previously mentioned (Brown et al., 2019; Geddie et al., 2001). As such, we may find that match between practice and interview context has more of an impact on the amount of information children say as opposed to how accurate they are.

Previous studies have found that introducing the ground rules can make children more cautious at responding to all questions, including ones they can answer (Gee et al., 1999; Nesbitt & Markham, 1999; Saywitz & Moan-Hardie, 1994). This type of over-cautious responding was not accounted for in this study as we did not code for children's spontaneous use of ground rules outside of the challenge questions. If this was the case, we might see that introducing the ground rules reduced the amount of information children gave when talking about the health and safety event. Consequently, future research could examine if different ground rule practice methods lead to differences in how much information children say when interviewed following a standard protocol.

Another possible explanation is that children might have interpreted the ground rules as concrete instructions and nothing further. When ground rules are being introduced, the child is instructed only to tell the truth and then tell the interviewer if they ask the child a question they do not know, understand or that are incorrect. Except for telling the truth, the ground rule instructions ask the child to monitor not what they say, but rather, what the interviewer says. As such, children may interpret these ground rules as only being a concrete instruction to be wary of the interviewer questions. If a child interprets the instructions this way, then they would not monitor the accuracy of their responses any more than what they would typically do. In turn, there would be no effect on general accuracy across the different practice conditions. Therefore, it is suggested that future research that examines accuracy with children's use of ground rules should account for concrete interpretations by including an additional instruction that asks that children think carefully about each question to determine if they can answer it or not. We would expect that this extra instruction would minimise the chance that children will only interpret the ground rule instructions in a literal manner.

Is Ground Rule Use Associated with the Accuracy of Details Children Report About a Recent Event?

To our surprise, there was no significant correlation between any of the rule types (or general rule use) and accuracy. This implies that children's ability to use ground rules in response to difficult questions has no relation to how accurately they may recall a recently experienced event. This suggests that practitioners should not assume that competency at using ground rules also indicates that a child will report less false details during the rest of the interview. So why might children's ability to detect and appropriately respond to difficult questions not also be associated with how accurate they are?

One assumption in ground rules literature is that some children may be better at applying ground rules because they have developed the ability to more accurately assess if they possess the knowledge being asked for or not (Brubacher et al., 2015). Therefore, if a child uses a ground rule when they are unsure how to respond, we could expect to see that children who used more ground rules to also report less false details. However, these results provide evidence that the task of answering difficult questions draws on different cognitive processes from those used in self-directed recall of an event that children remember well. This could be because when children are asked to freely recall an event, they are likely to search their memories and report what they can remember, generating further recall cues from the information shared. In contrast, ground rule challenge questions are initiated by the interviewer and purposefully tax either their memory or language processes. This is because children are asked about information, they do not have access to, is wrong, or the language or the syntax was challenging. As this is the first study to explore this topic, it is recommended that future research continue to explore the broader impacts of ground rule use on child testimony.

Additionally, the same explanations given in relation to the impact of practice method on accuracy also apply to this research question too. We might not have seen an association between overall accuracy and rule use because the children could have been editing their narratives for accuracy before verbalising them, the interview quality was good, children recalled the event well and so we saw ceiling effects in their accuracy, or it may depend on the type of question being asked.

Limitations

As already mentioned, the main limitation of this research project is the small sample size. If children are already competent at answering questions they do not know or that are incorrect, then we may not have had enough children in the lower age groups for a significant

variation to emerge between the groups. It is also recommended that future research in ground rules, especially when concerning the IDK and CM rules, should ensure adequate sample size and if possible, include children of preschool ages also.

Another limitation is that the coders of this study's data were not blind to the research questions being focused on this study. The coders were primarily processing the data for use in their respective research projects, which formed part of the wider research programme. However, the coders were blind to identifying information and details pertaining to the child such as their age when coding from the transcripts.

The key independent variable that was the focus of this study was the match between the practice and interview context; however, the author acknowledges that as the match increased, so did the intensity of practice. Uncoupling the intensity of practice from the match between practice and interview contexts would require adding additional conditions to the study. In these new conditions, children could practice ground rules to the same intensity levels as the existing conditions but with no changes in the variation of question type or practice within the practice narrative phase. It was not feasible to include these extra conditions in this study, but the author encourages future research which examines the effect on ground rules of the match between practice and interview context to consider how the intensity of practice may be accounted for within the research design.

Conclusion

This study was designed to bring together prior research on the use of ground rules in forensic interviews and Cognitive Learning to explore a possible new way of practising the ground rules. We assessed how increasing the match between the practice and interview contexts may support children to use the ground rules, 'I don't know', 'I don't understand' and 'Correct me' in response to difficult questions. We also accounted for broader effects of practice method and rule use on how accurately children report about a recently experienced event in an interview following the NICHD protocol (Lamb et al., 2018).

We found that amongst the conditions that included an element of practice, increasing the match between the practice and interview contexts does not significantly facilitate ground rule use. However, children were significantly better at answering questions they did not understand when they received any form of ground rule practice compared to not being introduced to ground rules at all. We also found that the accuracy of children's accounts were not affected by progressively aligning the match between the ground rules practice and interview contexts. Furthermore, there was no significant relationship between children's ability to apply ground

rules and how accurately they recalled the target event. These results could be due to ceiling effects in performance on dependent variables, the conditions of the event and interview or the use of different cognitive processes involved in applying different ground rules. The author proposes that future research should continue to explore under what conditions ground rules practice benefits children in forensic interviews and how this affects broader child testimony measures.

References

- Anderson, J. (2013). The CornerHouse Forensic Interview Protocol: An evolution in practice for almost 25 years. *APSAC Advisor*, 4, 2-7.
https://static1.squarespace.com/static/5da61df2de233e586746cdfe/t/5e3da17caafa9459cf/d2ad23/1581097340449/Anderson_2013_CornerHouse_Forensic_Interview_Protocol+%282%29.pdf
- APSAC Taskforce. (2012). Forensic Interviewing in Cases of Suspected Child Abuse. *The American Professional Society on the Abuse of Children (APSAC)*.
<https://www.apsac.org/guidelines>
- Braithwaite, D.W., & Goldstone, R.L. (2015). Effects of variation and prior knowledge on abstract concept learning. *Cognition and Instruction*, 33(3), 226-256.
<https://doi.org/10.1080/07370008.2015.1067215>
- Brown, D.A., Lamb, M.E., Lewis, C., Pipe, M., Orbach, Y., & Wolfman, M. (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., Gwynne, J., Kitto, O., & Stairmand, M. (2019). Developmental differences in children's learning and use of forensic ground rules during an interview about an experienced event. *Developmental Psychology*, 55(8), 1626-1639.
<https://doi.org/10.1037/dev0000756>
- Brubacher, S.P., Poole, D.A., Dickinson, J.J. (2015). The use of ground rules in investigative interviews with children: A synthesis and call for research. *Developmental Psychology*, 36(1), 15-33. <https://doi.org/10.1016/j.dr.2015.01.001>
- Brubacher, S.P., Roberts., K.P., & Powell, M. (2011). Effects of practicing episodic versus scripted recall on children's subsequent narratives of a repeated event. *Psychology, Public Policy, and Law*, 17(2), 286-314. <https://doi.org/10.1037/a0022793>
- Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development*, 11(3), 356-373.
<https://doi.org/10.1080/15248371003700015>
- Cordón, I.M., Saetermoe, C.L., & Goodman, G.S. (2005). Facilitating children's accurate responses: Conversational rules and interview style. *Applied Cognitive Psychology*, 19(3), 249-266. <https://doi.org/10.1002/acp.1090>

- Danby, M.C., Brubacher, S.J., Sharman, S.J., & Powell, M.B. (2015). The effects of practice on children's ability to apply ground rules in a narrative interview. *Behavioral Sciences and the Law*, 33(4), 446-458. <https://doi.org/10.1002/bsl.2194>
- Day, S.B., & Goldstone, R.L. (2012). The import of knowledge export: Connecting findings and theories for transfer of learning. *Educational Psychologist*, 47(3), 153-176. <https://doi.org/10.1080/00461520.2012.696438>
- Dickinson, J.J., Brubacher, S.P., & Poole, D.A. (2015). Children's performance on ground rules questions: Implications for forensic interviewing. *Law and Human Behavior*, 39(1), 87-97. <https://doi.org/10.1037/lhb0000119>
- Earhart, B., La Rooy, D.J., Brubacher, S.P., & Lamb, M.E. (2014). An examination of "Don't know" responses in forensic interviews with children. *Behavioral Sciences and the Law*, 32(6), 746-761. <https://doi.org/10.1002/bsl.2141>
- Ellis, L., Powell, M., Thomson, D., & Jones, C. (2003). Do simple "ground rules" reduce preschoolers' suggestibility about experienced and non experienced events? *Psychiatry, Psychology and Law: An Interdisciplinary Journal of Australia and New Zealand Association of Psychiatry, Psychology and Law*, 10(2), 334-345. <http://dro.deakin.edu.au/view/DU:30002231>
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics*. Sage Publications.
- Fivush, R., Haden, C.A., & Reese, E. (2006). Elaborating on elaborations: Role of maternal reminiscing style in cognitive and socioemotional development. *Child Development*, 77(6), 1568-1588. <https://doi.org/10.1111/j.1467-8624.2006.00960.x>
- Geddie, L.F., Beer, J., Bartosik, S., & Wuensch, K.L. (2001). The relationship between interview characteristics and accuracy of recall in young children: Do individual differences matter? *Child Maltreatment*, 6(1), 58-68. <https://doi.org/10.1177/1077559501006001006>
- Gee, S., Gregory, M., & Pipe, M. (1999). 'What colour is your pet dinosaur?' The impact of pre-interview training and question type on children's answers. *Legal and Criminological Psychology*, 4(1), 111-128. <https://doi.org/10.1348/135532599167716>
- Gelman, S.A., Raman, L., & Gentner, D. (2009). Effects of language and similarity on comparison processing. *Language, Learning and Development*, 5(3), 147-171. <https://doi.org/10.1080/15475440902824079>
- Gentner, D. (1983). Structure-mapping. A theoretical framework for analogy. *Cognitive Science*, 7(2), 155-170. [https://doi.org/10.1016/S0364-0213\(83\)80009-3](https://doi.org/10.1016/S0364-0213(83)80009-3)

- Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development*, 8(3), 285-307.
<https://doi.org/10.1080/15248370701446434>
- Höhle, B. (2009). Bootstrapping mechanisms in first language acquisition. *Linguistics*, 47(2), 359-382. <https://doi.org/10.1515/LING.2009.013>
- Hughes, M., & Grieve, R. (1980). On asking children bizarre questions. *First Language*, 1(2), 149-160. <https://doi.org/10.1177/014272378000100205>
- Kensinger, E.A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences*, 101(9), 3310-3315. <https://doi.org/10.1073/pnas.0306408101>
- Lamb, M.E., & Brown, D. A. (2006). Conversational apprentices: Helping children become competent informants about their own experiences. *British Journal of Developmental Psychology*, 24(1), 215-234. <https://doi.org/10.1348/026151005X57657>
- Lamb, M.E., Brown, D.A., Hershkowitz, I., Orbach, Y., & Esplin, P.W. (2018). *Tell me what happened* (2nd ed.). Wiley Blackwell
- Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. <https://doi.org/10.2307/2529310>
- Ministry of Justice. (2011, March). Achieving best evidence in criminal proceedings: Guidance on interviewing victims and witnesses, and guidance on using special measures.
https://www.cps.gov.uk/sites/default/files/documents/legal_guidance/best_evidence_in_criminal_proceedings.pdf
- Mulder, M.R., & Vrij, A. (1996). Explaining conversation rules to children: An intervention study to facilitate children's accurate responses. *Child Abuse & Neglect*, 20(7), 623-631.
[https://doi.org/10.1016/0145-2134\(96\)00050-6](https://doi.org/10.1016/0145-2134(96)00050-6)
- Nelson, K., Fivush, R. (2004). The emergence of autobiographical memory: A social cultural development theory. *Psychological Review*, 111(2), 486-511.
<https://doi.org/10.1037/0033-295X.111.2.486>
- Nesbitt, M., & Markham, R. (1999). Improving young children's accuracy of recall for an eyewitness event. *Journal of Applied Developmental Psychology*, 20(3), 449-459.
[https://doi.org/10.1016/S0193-3973\(99\)00027-1](https://doi.org/10.1016/S0193-3973(99)00027-1)
- Peters, W.W., & Nunez, M. (1999). Complex language and comprehension monitoring: Teaching child witnesses to recognize linguistic confusion. *Journal of Applied Psychology*, 84(5), 661-669. <https://doi.org/10.1037/0021-9010.84.5.661>

- Pratt, C. (1990). On asking children – and adults – bizarre questions. *First Language*, 10, 167-175. <https://doi.org/10.1177/014272379001002905>
- Price, H.L., Roberts, K.P., & Collins, A. (2013). The quality of children's allegations of abuse in investigative interviews containing practice narratives. *Journal of Applied Research in Memory and Cognition*, 2(1), 1-6. <https://doi.org/10.1016/j.jarmac.2012.03.001>
- Roberts, K.P., Brubacher, S.P., Powell, M.B., & Price, H.L. (2011). Practice narratives. In M. E. Lamb, D.J. La Rooy, L.C. Malloy, & C. Katz (Eds.), *Children's Testimony* (2nd ed., pp. 129-145). John Wiley & Sons.
- Salmon, K., & Reese, E. (2015). Talking (or not talking) about the past: The influence of parent-child conversation about negative experiences on children's memories. *Applied Cognitive Psychology*, 29(6), 791-801. <https://doi.org/10.1002/acp.3186>
- Saywitz, K., & Moan-Hardie, S. (1994). Reducing the potential for distortion of childhood memories. *Consciousness and Cognition*, 3(3), 408-425. <https://doi.org/10.1006/ccog.1994.1023>
- Saywitz, K.J., Nathanson, R., & Snyder, L.E. (1999). Facilitating the communicative competence of the child witness. *Applied Developmental Science*, 3(1), 58-68. https://doi.org/10.1207/s1532480xads0301_7
- Schmidt, A.F., & Finan, C. (2018). Linear regression and the normality assumption. *Journal of Clinical Epidemiology*, (48), 146-151. <https://doi.org/10.1016/j.jclinepi.2017.12.006>
- Schneider, W. (2002). Memory development in childhood. In U. Goswami (Eds.), *Blackwell Handbook of Childhood Cognitive Development* (pp. 236-256). Blackwell.
- Vogl, S. (2012). Children between the age of 5 and 11: what "don't know" answers tell us. *Quality & Quantity*, 46(4), 993-1011. <https://doi.org/10.1007/s11135-011-9438-9>
- Waterman, A.H., Blades, M., & Spencer, C. (2000). Do children try to answer nonsensical questions? *British Journal of Developmental Psychology*, 18(2), 211-225. <https://doi.org/10.1348/026151000165652>
- Waterman, A.H., & Blades, M. (2011). Helping children correctly say "I don't know" to unanswerable questions. *Journal of Experimental Psychology*, 17(4), 396-405. <https://doi.org/10.1037/a0026150>
- Westera, N., Zajac, R., & Brown, D. A. (2016). Witness interviewing practices in New Zealand. In D. Walsh, G.Oxburgh, A.Redlich, & T.Myklebust (Eds.), *International developments and practices in investigative interviewing and interrogation. Volume 1: Victims and witnesses* (pp. 87-98). Routledge.

Appendix A

Health and Safety Event Script

1. Health and Safety Event Leader and Interrupter

‘Hi everyone, my name is Welcome to our health and safety event. I’m here today with some of my helpers to talk to you about how to keep healthy and safe. There are four stations where you will learn how to look after yourselves if you get sick or hurt, and how to stop this from happening. You will all have a turn at two of the stations.’

‘At the “Heartbeats” station the leader is _____.’

‘At the “Temperatures” station the leader is _____.’

‘At the “Care of cuts” station the leader is _____.’

‘At the “Hazards” station the leader is _____.’

‘At one of the stations you visit, you will do activities with a partner.’

‘So first we will split up into teams. Stay in your teams as you go around the stations.’

‘Alright time to get started. I’ll read your names out and tell you which station to go to.’

‘The yellow team is [names] you are going to the _____ station first’

‘The red team is [names] you are going to the _____ station first’

‘The green team is [names] you are going to the _____ station first’

‘The orange team is [names] you are going to the _____ station first’

[Event Leader reads out the teams of children and sends them to stand by their Station Leader]

[Event Leader helps any station leaders needing help and records on event record form].

[Give a two-minute warning when appropriate]

[When all groups are finished and the station leaders have reset, the Event Leader will instruct each group to go to their next station]

[After the children have been sent to their second activity the interrupter will enter the room]

[Interrupter bursts in (acting panicked) and rushes around each station asking for the Event Leader]

Station Leaders to Group: “Hey, we’d better just stop for a minute and see what is going on.”

Interrupter to Station Leaders: “Hey, I need to find [Event Leader name], do you know where she/he is?”

Station Leaders to Interrupter: “No, sorry, try over there.”

[After asking for the Event Leader at the last (4th) station the interrupter finds the Event Leader, and they meet in the middle of the room]

Interrupter to Event Leader: “Here you are! You can’t have these things now; I’m supposed to be giving a demonstration in Miramar in an hour.”

Event Leader to Interrupter: “But I’ve had this presentation scheduled for weeks.”

Interrupter to Event Leader: “Well, it wasn’t in the diary. I must leave now to be in Miramar in time. What are you going to do?”

Event Leader to Interrupter: “Me? It’s not my fault! I had it up on the board.”

Interrupter to Event Leader: “Yes, well, we don’t have time to argue about that. I need to leave now to get there in time. How long will you be?”

Event Leader to Interrupter: “I’ll be another 15-20 minutes by the time I pack up. But look, *[gesture toward the spare equipment at stations]* I have heaps of spare stethoscopes, we don’t need all the pictures, and there’s an extra copy of the cuts presentation on the USB, and we have some spare thermometers. Why don’t you take those and then both of us are okay?”

Interrupter to Event Leader: “Good idea. Thanks a lot _____. Sorry, I burst in I was in a panic.”

Event Leader to Interrupter: “No problem.”

[Leader and interrupter walk around the teams, collecting extra equipment. Interrupter apologises for interrupting]

Interrupter to Station Leaders: “Sorry for bursting in like that, could I please have the spare stuff you’re not using.”

Station Leaders to Interrupter: “Sure, here you go” *[Station Leader hands over spare equipment]*

Event Leader to Interrupter: “Good luck in Miramar; travel safe.”

‘Great work today everyone, we have learned a lot today about health and safety. We hope you have had fun. To say thank you, we have a small prize for everyone. Please put this in your school bag when you go back to class so that the other children don’t ask you questions about it.’

2. Care of Cuts Station

‘Hey guys, my name is _____, and now we are at the Care of Cuts” Station’.

[Hand out name tags if this is the first group]

[Use slideshow presentation to work through station activities]

‘Now we are going to have a chat about how we can look after ourselves if we get a cut.

First, we are going to watch a short video about Arthur who has gone to get a tyre from the dump. Now, I want you all to watch really, really carefully and pay attention to everything that is happening in the video.’

[Play Arthur video clip]

[After the video discuss with the children about the below questions]

- What happened to Arthur?
- How did Arthur cut his knee?
- Was it a small cut or a big cut?
- If we were there, how could we have helped Arthur?

[Continue to work through slideshow]

‘Ok now everyone hold out your left pointer finger out.’ *[show them]*

‘I am going to draw a pretend cut on your finger.’

[Draw pretend cut on each child’s left pointer finger]

‘So, when we have a cut, the first thing we need to do is put pressure on the cut to stop the bleeding. Se everyone put pressure on your cut!’

‘The second thing we do is raise the cut up above your heart.’

‘Now everyone take an antiseptic wipe *[hand out one wipe to each child]* and wipe your cut to stop it from getting infected.’

‘Next, you cover the cut with a plaster, *[hand out one plaster to each child]* so I will give you all a plaster, and you can do this last step. You all looked after your cuts so well!’

So, what were the steps for caring for a cut again? *[Go through the points on the slide]*. These are really important steps for looking after a small cut and these could also be used to help Arthur with his cut.’

‘If it’s a big cut that is bleeding a lot, then you should always get an adult to help you.

Sometimes big cuts need to get stitches to make them all better. And always make sure that any blood is cleaned up.’

Right now, show me all your plasters, and I’ll take a photo!’

[Take photo of children holding their pretend cuts in the air]

[If the first group] ‘Great work everyone, we’re all finished at this station. While we wait for everyone to finish up, I want you to all go and sit in the centre and wait for our Event Leader ____ to tell you what station you will go to next. Remember, some groups may still be working hard, so let’s move and wait quietly, okay? Great, you may go now.

[If the last group] ‘Well done we’re finished. Thank you so much for coming and talking to us. I hope you had fun and learn a lot. Because you all worked so well the Event Leader has a thank-you prize for you, so everyone go back to the middle and wait quietly for the Event Leader to tell you what to do now.’

3. Temperature Station

‘Hey guys, my name is _____, and now we are at the “TEMPERATURES” Station.’

[Hand out name tags if this is the first group]

‘At the temperatures station, we are going to learn about how we take our temperature. Our temperature is how hot or cold our body is. It is important to check our temperature when we think we are sick. When we are sick, our body often heats up to fight off bugs.’

‘First, I need you to get into pairs.’

‘Now in your pairs decide who is the number one partner, you will go first, and who is the number two partners, you will go second.’

‘Alright, put your hands up number ones’

‘Now hands up number twos.’

‘Sometimes your mums and dads might check your temperature on your forehead, like this.’

[Station Leader puts back of their hand to their forehead]

‘Now, Number ONES I want you to check your partner’s temperature on their forehead. How does it feel?’

‘Now Number TWOS, you can have a go now. Put the back of your hand to your partner’s forehead. How does it feel?’

‘Another way we can check our temperature is using a thermometer.

‘This is a thermometer *[hold up a skinny thermometer]*. Now say it with me just so I know how clever you all are ... what is this called? A THERMOMETER. The temperature the thermometer tells can be different when we measure temperature in different places.’

‘We are going to practice using this skinny thermometer, so Number ONES all take a thermometer’ *[turn on and hand out one skinny thermometer to each pair]*.

‘Now, Number ONES, take your partner’s temperature under their left arm. So all Number TWOS put your left arm out, and number ONES pop the thermometer under their arm and get them to put their arm down.

We will leave it there for 10 seconds, so everyone count to ten with me 1, 2, 3, 4.....’

‘Ok Number ONES, now we are going to check our partner’s temperature behind their left knee. This isn’t usually where we take our temperature, and we will see if the temperature is different. So, all number ONES put the thermometer behind your partner’s left knee and get them to close their leg around it.

Ok everyone count to ten. Is it the same or different to the other temperature?’

‘Great job! Now Number ONES take an antiseptic wipe *[hand out one antiseptic wipe to each pair]*. Wipe the thermometer down *[have a rubbish bag ready for dirty wipes]*.

Ok great now Number TWOS take the thermometer’.

‘Put the thermometer under your left partner’s arm. Now everyone count to ten. Check the temperature.

Now put it behind your partner’s left knee. Everyone count to ten. Great, now check the temperature, is it the same or different.’

‘It’s okay if you get a LOW reading it just means that it is too difficult for the thermometer to get a reading there, so it is much better taking your temp under your arm.’

‘Another place we can check our temperature is in our ear. This gives us the best reading.’

[Hand out one ear thermometer, a temperature record sheet and pencil to each pair]

Number ONES put the thermometer gently in your partner’s left ear and press the button.

Number TWOS have a look at the number on the thermometer screen and circle the number on the sheet that matches. I will come around and have a look and help you do this.

Number ONES wipe down the thermometer with this wipe [hand out an antibacterial wipe to Number ONES] while your partner is filling out the sheet.

‘Number TWOS you take the thermometer and gently put it in your partner’s left ear and press the button. Now,

Number ONES you circle the number the matches the number on the screen’ *[Help or have children write the whole number out, e.g. 36.65 as they often have the same whole number]*.

Number TWOS wipe down the thermometer with this wipe [hand out an antibacterial wipe to Number TWOS] while your partner is filling out the sheet.

‘Great work, everyone, were you hot or cold. Really? Great!’

[If this is the first activity for the group] ‘Great work everyone, we’re all finished at this station.

While we wait for everyone to finish up, I want you to all go and sit in the centre and wait for our Event Leader ____ to tell you what station you will go to next. Remember, some groups may still be working hard, so let’s move and wait quietly, okay? Great, you may go now.

[If this is the last activity for the group] ‘Well done we’re finished. Thank you so much for coming and talking to us. I hope you had fun and learn a lot. Because you all worked so well the Event Leader has a thank-you prize for you, so everyone go back to the middle and wait quietly for the Event Leader to tell you what to do now.’

4. Hazards Station

‘Hey guys, my name is _____, and now we are at the Hazard Station.’

[If the first group, hand out name tags]

‘Now we are going to learn about hazards and how to make dangerous or unsafe things, safer.’

‘We are going to look at some pictures that show some hazards that aren’t very safe.’

[There is one example picture, MUST complete pictures 1-5 with all teams, pictures 6-10 are additional, complete as many as there is time for]

‘When you are looking at the pictures, I want you all to look carefully and think about:

1. What is unsafe in the picture
2. WHY *[emphasise]* it is unsafe and
3. HOW we can try and fix it.’

‘Let’s practice with the first one together. Please put your hand up when you have an answer.’

[Show example kitchen picture, restate instructions if necessary]

- ‘What do you think is unsafe in this picture?’ *[choose one child to answer]*
- ‘Why is it unsafe?’ *[choose a different child to answer]*
- ‘How could we make it less dangerous?’ *[choose a different child to answer]*

[If the first group] ‘Great work everyone, we’re all finished at this station. While we wait for everyone to finish up, I want you to all go and sit in the centre and wait for our Event Leader ____ to tell you what station you will go to next. Remember, some groups may still be working hard, so let’s move and wait quietly, okay? Great, you may go now.’

[If the last group] ‘Well done we’re finished. Thank you so much for coming and talking to us. I hope you had fun and learn a lot. Because you all worked so well the Event Leader has a thank-you prize for you, so everyone go back to the middle and wait quietly for the Event Leader to tell you what to do now.’

5. Heartbeats Station

‘Hey guys, my name is _____, and now we are at the Heartbeat Station.’

[If the first group, hand out name tags]

‘Now at the heartbeats station, we are going to learn about how to measure our heartbeat.

Doctors will often check our heartbeat because if our heart is beating too fast or too slow, it can tell the Doctor that we are not feeling well. ‘

‘This is called a stethoscope *[hold up stethoscope]*, and Doctors use this to listen to our heart.’

‘So, what was this instrument called again? Say it with me... A STETHOSCOPE!’

‘We are going to practice using these, so I would like you all to partner up and decide if you are the number one partner or the number two partner.’

‘Number one's hands up – you are going first.’

‘The number one partner will use the stethoscope first and go first the whole time.’

‘Number twos hands up – you will go second.’

‘So again, hands up number one’s, and hands up number twos.’

‘Ok now hands up number ONES, I want you to take a stethoscope.

All the number ONES put the bits in your ears and tap the round bit at the end. If you don’t hear the tap, twist the end around like this and see if the tapping is louder *[show them]*.

Alright, all the number ONES get your stethoscopes ready.’

‘I want you to listen to your partner’s heartbeat on the left of their chest. Do it all together. Did you hear it? This is where Doctors usually listen to our hearts.’

‘They don’t usually listen to our stomachs, though. So, number ONES I now want you to listen to your partner’s stomach. Did any of you hear anything? *[NO?!]* That’s because this isn’t a very good place to listen to our partner’s heartbeat.’

‘Okay, now number ONES take an antiseptic wipe and I want you to wipe the earpieces *[hand out wipes]*.

Alright, number TWOS, you take the stethoscope.’

‘First, listen to the left of your partner’s chest.’

‘Now listen to your partner’s stomach. Did you guys hear anything there? No? See it’s not a very good place to check for a heartbeat, is it?’

‘Okay, now number TWOS take an antiseptic wipe and I want you to wipe the earpieces *[hand out wipes]*

‘Another way of listening to our heartbeat is by feeling our pulse on our wrist – *a pulse is when we feel the heartbeat rather than hearing it.*’

‘We do this by pressing our two fingers down on our wrist *[show them]*.

So, number ONES you go first.’

‘Press two fingers down on your partner’s wrist like this *[show them on yourself]*. Do you feel anything? Swap over now and let number TWOS have a go. Did you feel it? This is how we normally check a pulse.’

‘Okay so now let’s see if we can check our partner’s pulse on their ankle. Number ONES press two fingers on your partner’s ankle *[show them on yourself]*. Did you feel anything?’

‘Now, number TWOS – you have a go. Press 2 fingers on your partner’s ankle. Did you feel anything? It is not a very good place to feel your pulse on your ankle, so this is why you probably didn’t feel anything here, and this is why we normally feel our pulse on our wrist.’

‘Right now, standing next to your partner, I am going to take a take a photo!’

[One partner to hold up stethoscope for photo]

[If the first group] ‘Great work everyone, we’re all finished at this station. While we wait for everyone to finish up, I want you to all go and sit in the centre and wait for our Event Leader ____ to tell you what station you will go to next. Remember, some groups may still be working hard, so let’s move and wait quietly, okay? Great, you may go now.

[If the last group] ‘Well done we’re finished. Thank you so much for coming and talking to us. I hope you had fun and learn a lot. Because you all worked so well the Event Leader has a thank-you prize for you, so everyone go back to the middle and wait quietly for the Event Leader to tell you what to do now.’

Appendix B

Challenge questions versions used in interviews

Version 'A' for children who attended Temperatures and Treatment of Cuts stations

Rule type	Question type		
	Open-ended	Directive (Wh-)	Option posing
'I don't know'	"Tell me everything that happened with the leaders before they came and got you?"	"What was the name of the girl who cut her finger in the slideshow?"	"The woman who came in, did she arrive by bicycle?"
'I don't understand'	"Tell me about measuring febrility?"	"When did the bellicose woman leave the [hall/classroom]?"	"Did you put your thermometer next to your tympanum?"
'Correct me'	"Earlier you said you did different activities, tell me about what you did at the [wrong station] hazards station?"	"Where did you draw the pretend cut?"	"Was your temperature the same [use higher or low if the temperature was the same] as your partners or [higher/lower – use incorrect]?"

Version 'B' for children who attended Temperatures and Treatment of Cuts stations

Rule type	Question type		
	Open-ended	Directive (Wh-)	Option posing
'I don't know'	"Tell me everything that happened with the leaders before they came and got you?"	"The woman who came in, what colour is her bicycle?"	"You know the girl in the slideshow who cut her finger, was her name Sarah?"
'I don't understand'	"Tell me about measuring febrility?"	"When did you put the thermometer next to your tympanum?"	"Did Arthur hurt his patella?"
'Correct me'	"Earlier you said you did different activities, tell me about what you did at the	"Where was the man who needed the spare equipment going?"	"Did you use a red pen to draw a cut on your knee?"

[wrong station] hazards
station?"

Version 'C' for children who attended Temperatures and Treatment of Cuts stations

Rule type	Question type		
	Open-ended	Directive (Wh-)	Option posing
'I don't know'	"Tell me everything that happened with the leaders before they came and got you?"	"What was the leader's temperature?"	"The woman who came in, did she arrive by bicycle?"
'I don't understand'	"Tell me about measuring febrility?"	"Which part of his crus did Arthur hurt?"	"Did the bellicose woman leave the [hall/classroom] after picking up spare equipment?"
'Correct me'	"Earlier you said you did different activities, tell me about what you did at the [wrong station] hazards station?"	"When the leader saw the temperature was 42 (wrong temperature), what did she say?"	"Was the woman who needed the spare equipment going to Karori?"

Version 'A' for children who attended Hazards and Heartbeat stations

Rule type	Question type		
	Open-ended	Directive (Wh-)	Option posing
'I don't know'	"Tell me everything that happened with the leaders before they came and got you?"	"In the picture of the girl with scissors, what happened after she ran with scissors?"	"The woman who came in, did she arrive by bicycle?"
'I don't understand'	"Tell me more about measuring arrhythmia?"	"When did the bellicose woman leave the [hall/classroom]?"	"Did someone auscultate your pulse?"

‘Correct me’	“Earlier you said you did different activities, tell me about what you did at the [wrong station] treatment of cuts station?”	“In the picture of the garden, who got hurt stepping on the rake?”	“Did you like your leader’s tiger stethoscope?”
--------------	---	--	---

Version ‘B’ for children who attended Hazards and Heartbeat stations

Rule type	Question type		
	Open-ended	Directive (Wh-)	Option posing
‘I don’t know’	“Tell me everything that happened with the leaders before they came and got you?”	“The woman who came in, what colour was her bicycle?”	“Did the girl who ran with scissors fall and cut her brother?”
‘I don’t understand’	“Tell me more about measuring arrhythmia?”	“Who auscultated your pulse?”	“In the picture of the road, was the man on the zebra crossing running with impurity?”
‘Correct me’	“Earlier you said you did different activities, tell me about what you did at the [wrong station] treatment of cuts station?”	“Where was the man who needed the spare equipment going?”	“In the picture of the garden, did the boy hurt his foot when he stepped on the rake?”

Version ‘C’ for children who attended Hazards and Heartbeat stations

Rule type	Question type		
	Open-ended	Directive (Wh-)	Option posing
‘I don’t know’	“Tell me everything that happened with the	“How much did the stethoscopes cost the	“The woman who came to take the spare

	leaders before they came and got you?"	leader at the heartbeat station?"	equipment, did she arrive by bicycle?"
'I don't understand'	"Tell me more about measuring arrhythmia?"	"Who auscultated your pulse?"	"In the picture of the road, was the man on the zebra crossing running with impigrity?"
'Correct me'	"Earlier you said you did different activities, tell me about what you did at the [wrong station] treatment of cuts station?"	"Where was the leader wearing her Bear stethoscope?"	"Was the woman who came in going to Karori?"

Appendix C

All ground rule practice questions by condition

Business as usual condition

Rule type	Question type		
	Open-ended	Directive (Wh-)	Option posing
'Tell the truth'	-	-	<p>"If I say that my shoes are red [or a colour they are not] is that true or not true?"</p> <p>"If I say that I am sitting down now, would that be true or not true?"</p>
'I don't know'	-	<p>"What is my dog's name?"</p> <p>"What were you doing on the 1st of June [use another date if this is their birthday] two years ago?"</p>	-
'I don't understand'	-	<p>"What are you most adroit at?"</p> <p>"Where is the querulous cat?"</p>	-
'Correct me'	-	<p>"If I said that you are a 2-year-old girl/boy [when interviewing a 5-year old boy, etc.] what would you say?"</p> <p>"If I said that you live in Australia, what would you say?"</p>	-

Note. All questions are delivered even if the child gets initial questions correct.

Extended Practice

Rule type	Question type		
	Open-ended	Directive (Wh-)	Option posing
'Tell the truth'	-	-	<p>"If I say that my shoes are red [or a colour they are not] is that true or not true?"</p> <p>"If I say that I am sitting down now, would that be true or not true?"</p> <p>"If I say that I am wearing a hat, is that true or not true?"</p>
'I don't know'	<p>"Tell me all about my last birthday party?"</p> <p>"Tell me all about the first time you ate with a spoon?"</p>	<p>"What is my dog's name?"</p> <p>"What were you doing on the 1st of June [use another date if this is their birthday] two years ago?" what would you say?"</p>	<p>"Did I go to the movies last night?"</p> <p>"Did I go dancing last night?"</p>
'I don't understand'	<p>"Tell me all about your curmudgeon teacher?"</p> <p>"Tell me all about how rats are erinaceous to hedgehogs?"</p>	<p>"What are you most adroit at?"</p> <p>"Where is the querulous cat?"</p>	<p>"Is my shirt gridelin?"</p> <p>"Did you uhtceare this morning?"</p>
'Correct me'	"Tell me about the Prime Minister's visit to your school this morning?"	"If I said that you are a 2-year-old girl/boy [when interviewing a 5-year-old boy, etc.] what would you say?"	"Is your hair pink or is your hair purple [colours their hair is not] today?"

“Tell me all about sleeping at school last night?”	“If I said that you live in Australia, what would you say?”	“Do you live in Christchurch or Dunedin?”
--	---	---

Note. Questions are presented in the order they were asked. If a child got the first question correct for the ‘I don’t’ know’, ‘I don’t understand’ and ‘Correct me’ rules correct, the second question was not asked.

Extended Practice and Practice Narrative (Version 1)

Rule type	Question type		
	Open-ended	Directive (Wh-)	Option posing
‘I don’t know’	“Tell me what was happening on your street before you woke up this morning?”	“Which one of your friends woke up the earliest this morning?”	“Did your teacher eat yogurt for breakfast this morning?”
‘I don’t understand’	“Tell me about making your bag pendant?”	“When did you feel the most acrasial this morning?”	“Is [just disclosed detail] a hemerine part of your morning?”
‘Correct me’	“Tell me about the hot air balloon that was in the sky on your way to school?”	“How long did it take you to go home and get the school bag you forgot this morning?”	“Did you leave the classroom when the fire engine came to school this morning?”

Note. Children were asked practice questions from the Extended Practice condition and the above questions embedded in the practice narrative.

Extended Practice and Practice Narrative (Version 2)

Rule type	Question type		
	Open-ended	Directive (Wh-)	Option posing
'I don't know'	"Who is your best friend? Tell me everything that happened when they got up this morning?"	"What did your teacher eat for breakfast this morning?"	"Were there people walking on your street before you woke up this morning?"
'I don't understand'	"Tell me about feeling acrasial this morning?"	"What are your hemerine morning activities?"	"Did you make your bag pendant?"
'Correct me'	"Tell me everything that happened when the fire engine came to school this morning?"	"What colour was the hot air balloon that was in the sky on your way to school/preschool?"	"Did you go back home to get the school bag you forgot this morning?"

Note. Children were asked practice questions from the Extended Practice condition and the above questions embedded in the practice narrative.

Extended Practice and Practice Narrative (Version 3)

Rule type	Question type		
	Open-ended	Directive (Wh-)	Option posing
'I don't know'	"Tell me everything that happened when your teacher had breakfast this morning?"	"How many people were walking on your street before you woke up this morning?"	"Did you wake up earlier than all your friends this morning?"
'I don't understand'	"Tell me everything about a hemerine part of your morning?"	"How did you make your bag pendant?"	"Did you feel acrasial when you woke up today?"
'Correct me'	"Tell me about going home to get your	"What did you do when the fire engine	"Was the hot air balloon in the sky on

school bag you forgot
this morning?"

came to school this
morning?"

your way to school
red and white?"

Note. Children were asked practice questions from the Extended Practice condition and the above questions embedded in the practice narrative.