“You can Point to This Card”: Trialling the use of Visual Aids in Teaching Children to use Ground Rules in Forensic Interviewing

By

Tui Shaw

A thesis

submitted to the Victoria University of Wellington

in fulfilment of the requirements for the degree of

Master of Science in Psychology

Victoria University of Wellington

2022

Abstract

Often when children are interviewed, they are taught ground rules to help guide their responses to questions. Ground rules are utilised frequently in forensic interviewing, but there remain questions about their functionality and efficacy. This research examined the use of visual cue cards to teach children ground rules, First, we piloted the use of realistic photographs versus cartoon-like emoticons when responding us ground rules, and children showed an overwhelming preference for emoticons, so visual cue cards were developed using emoticons displaying gestures representing the ground rules. Then, 89 children aged 5-12 years participated in a staged event and were subsequently interviewed about it. Children were randomly assigned to a "business as usual" (BAU) control group (verbal instruction only) or to the visual aids (VA) group (verbal and visual instruction). We predicted that children in VA would respond better to training questions, and to challenge questions; and that they would spontaneously utilise ground rules more than those in BAU. Results did not support these hypotheses. Performance in responding with ground rules was relatively poor despite condition, with no significant differences in responding observed. Spontaneous use of ground rules was negatively correlated with age, indicating that older children spontaneously utilised ground rules significantly less than younger children. Theoretical research implications are discussed. We conclude that the current method of training children to use ground rules are not effective, and we have demonstrated that visual cue cards do not provide sufficient support for children in this context. Future research should explore a more comprehensive training method which includes an extended verbal instruction of the ground rules and incorporates more varied and nuanced opportunities for children to practice using ground rules.

Acknowledgements

To my supervisor, Dr Deirdre Brown, thank you for all the guidance, support, and mentoring that you have offered me throughout this journey. I admire your knowledge, your kindness, your flexibility, your approachability, and most importantly your patience with me as I was grappling with the unkind world of statistical analysis.

To my family and friends, thank you for putting up with my rants and ravings and for supporting me as a student all these years. You have all consistently shown me unconditional love and support.

To my husband Nick, thank you for everything that you do for me. From the endless proofreading of multiple drafts, to encouraging me to take a break when I needed it, but also for gently pushing me to power on when I was getting stuck. I truly don't think I could have done this without.

To all the members of the Applied Developmental Psychology lab (whether you be fellow students, volunteers, or paid employees), thank you for your support in the lab. Sharing a space with such wonderful people through this journey certainly helped to ease the load and discuss ideas openly.

And lastly, thank you to each and everyone of the many children who participated in this research – I literally could not have done this without you! Your little faces and funny conversations during our interviews will be fondly remembered for many years to come.

Table of Contents

List of Tables, Figures, and Appendices

"You can Point to this Card": Trialling the use of Visual Aids in Teaching Children to use Ground Rules in Forensic Interviewing

How do children develop their ability to understand and answer questions? Children are often asked important questions, for example at the doctors, by teachers, in research, or in clinical and legal contexts. From a young age children learn that they should always answer adults' questions (Nelson, 2013) even when they do not know the answer (Waterman et al., 2004) or they do not understand the question (Lamb & Brown, 2006). In high stakes situations such as forensic interviews, where children are questioned about allegations of abuse, the answers given by children can critically affect subsequent decisions made by adults (Lamb et al., 2018). It is crucial that children do not answer inappropriately because it may undermine the forensic investigation, the credibility of their evidence, and the verdicts reached. Hence, it is very important to gain a thorough understanding of how to support children to navigate the challenges of a forensic interview, especially in New Zealand given our alarmingly high rates of child abuse within the developed world and the high frequency of forensic interviews being conducted with vulnerable children (de Haan et al., 2019).

Interview Conversations are Different to Normal Conversations

In everyday conversation, children do not interact with adults in the same way as they are expected to during a forensic interview. The nature of formal interviewing conversations between adults and children, where the child is the expert and the adult is attempting to gather information, is a unique and often confronting experience for the child (Lamb et al., 2018). In daily conversations there is often no consequence attached to errors in children's descriptions of their experiences – accuracy does not hold the same premium that it does in a forensic interview (Golding et al., 2015). Children have also been socialised from a very young age to understand certain conversational conventions. For example, children are often encouraged to guess answers to adults' questions as the adult often assumes the role of an

expert or teacher in these interactions (Lamb & Brown, 2006; Zajac & Brown, 2018). Children may have been discouraged from commenting on or correcting adults' mistakes based on social conventions around respecting one's elders. Children learn to make guesses about what a question (or statement) means, even if they do not understand the entirety of it.

Ground Rules as a Supportive Interviewing Practice

To help children recognise that adults have different expectations of them in a forensic setting, interviewers often teach them 'ground rules'. These ground rules are guidelines which scaffold the child's conversational style during an interview, theoretically allowing them to understand what is expected of them and how they should respond to questions (Brubacher et al., 2015). For example, to say "I don't know" when they are unsure of the answer to a question; "I don't understand" when they are unsure what the interviewer means; or "that's not right" when the interviewer makes a mistake (Lamb et al., 2018). Ground rules are designed to help children learn that sometimes an interviewer may ask them incorrect or difficult questions, and since they are the experts of their experiences, it is up to them to let the interviewer know when they cannot answer a question (Lyon, 2014).

Ground rules have been included in wide variety of interviewing protocols dating back to the 1990s (Dickinson et al., 2015). In New Zealand, specialist interviewers are trained to evaluate allegations of child abuse (Westera et al., 2016) with a protocol that is closely modelled on the National Instititutes of Child Health and Human Development (NICHD) Investigative Interview protocol (Lamb et al., 2018). Ground rules are introduced to children at the outset of the interview and practiced. However, despite widespread use of ground rules in forensic interviewing, we know little about how well children of different ages understand them, what impact they have on children's behaviour, and the most appropriate approach for teaching children how to use these ground rules effectively (Brubacher et al., 2015).

Evidence for the Utility of various Ground Rules in Forensic Interviewing

Ground rules are commonly used in forensic interviewing in New Zealand and worldwide (Dickinson et al., 2015) so it is important to understand the rationale and the literature base behind each of the three ground rules being examined in this research. The purpose of the "I don't know" rule is to encourage children to let the interviewer know when they are not sure of the answer to their question. The "I don't know" rule is the most studied of all the ground rules (Brubacher et al., 2015). Explicitly teaching children to say "I don't know" when they are unsure of the answer significantly reduces the inaccuracies of children's reports, especially when explanation of the rule goes beyond a simple instruction and is accompanied by practice using the rule, feedback, and reinforcement about the child's performance using the rule (Saywitz & Moan-Hardie, 1994; Nesbitt & Markham, 1999; Dickinson et al., 2015; Peterson & Grant, 2001).

For example, studies that have trained children extensively about how, when, and why to use the "I don't know" rule, alongside feedback, reinforcement, and practice, increased children's use of the rule and the accuracy of their responses to challenging questions which had been specifically designed to test their use of the ground rule (Saywitz & Moan-Hardie,1994; Nesbitt & Markham, 1999; Dickinson et al., 2015). In other research, Peterson and Grant (2001) provided a simple rule statement to children in their study and did not find positive impacts on children's accuracy. Moreover, in a recent analysis of transcripts of interviews with children about alleged abuse, simple rule statements were not sufficient to increase children's likelihood of telling the interviewer that they were unsure of an answer (Earhart et al., 2014).

Children are taught to say "I don't understand" because interviewers may refer to concepts or use language and words within the forensic context that are unfamiliar to children (Cooper et al., 2010). This was first demonstrated in Hughes and Grieve's (1980) seminal

research. Children were asked non-sensical questions such as “Is milk bigger than water?”. Hughes and Grieve reported that despite the bizarreness of the questions, children often attempted to answer them. Similarly, Pratt (1990) demonstrated that children would answer a question even though they had shown that they knew that the question did not make sense. Saywitz et al. (1999) examined the efficacy of instructing children to say “I don’t understand”. Children who received training, practice, feedback, and reinforcement for indicating that they did not understand a question were significantly more accurate in their accounts of what had happened when questioned about a previously staged event than children whom were instructed simply to let the interviewer know if they would like them to rephrase the question.

The third ground rule under examination is the “that’s wrong” rule. The purpose of this ground rule is to encourage children to let the interviewer know when they have said something that is incorrect, and to provide them with a correction. Research has found that interviewers often make mistakes or conflate information when summarising children’s statements, and that children are not likely to correct these mistakes (Roberts & Lamb, 1999).

As with the “I don’t know” rule, when the “that’s wrong” rule is taught to children thoroughly, including an explanation for the rational of the rule, practice using it, and feedback on their performance; it has been shown to decrease children’s inaccuracies and increase their likeliness of telling the interviewer when they have said something that is incorrect (Geddie et al., 2001; Gee et al., 1999; Saywitz & Moan-Hardie, 1994). However, when the “that’s wrong” rule is delivered to children as a simple instruction statement and there is no practice involved (e.g. Ellis et al., 2003); or when the rule training is delivered by a separate person than the interviewer (e.g. Geddie et al., 2001), the benefits of using the “that’s wrong” rule are not observed. As with the “I don’t understand” rule, this further

emphasizes the importance of the way children are trained to use ground rules in forensic interviews.

Clearly then, ground rules may be beneficial for children and aid accurate reporting, but the way in which they are introduced and taught to children has an impact on whether children are likely to benefit from using them. Training and practice with using the rules is important for children to gain any benefit from it.

Developmental Changes in the use of Ground Rules

It is reported that children are still developing many of the capabilities required to provide a coherent testimony, and despite this, pressures beyond their capacities are still placed on them during forensic interviews, resulting in a mismatch between the capabilities of children and the requirements of the system (Saywitz, 2002). When developmental differences are not accounted for, simple questions that a 10-year-old child may find easy could be met with confusion and misunderstanding by a 4-year-old, for example "how many times did it happen?" (Saywitz, 2002). These developmental differences have called for the inclusion of ground rules in forensic interviews to help support children to answer tricky questions which they may still be otherwise developing the skills to answer appropriately.

However, there has been a consistent lack of acknowledgement of developmental influences and how these may impact children's understanding of, and ability to, apply ground rules. Brubacher et al. (2015) conducted an analysis into the evidence basis for the use of each of the identified ground rules and suggested that because social and cognitive abilities are continually developing throughout childhood, and at differing rates among children, this could impact how children interpret, understand, and apply the rules throughout an interview.

Aguiar et al. (2012) investigated children's ability to identify gaps in their knowledge, and to fill these gaps by seeking information from an appropriate source. This is similar to

what children are expected to do when utilising the "I don't know" rule – they are required to identify that they do not know the answer to a question or how to respond to a statement, and to make this known to the interviewer. Aguiar et al. found that while 4-5 year olds were skilled at assigning correct experts to domains of questions (e.g. a doctor is the person to ask a question regarding medicine), they were not able to identify that they did not actually know the answers to the questions. However, 6-year-old children were able to perform this task correctly and identify that they did not know the answers, as well as assign an appropriate expert to seek knowledge from. This suggests that younger children will not be able to correctly use the "I don't know" rule, while children aged 6 and over will be able to use it as they are likely to have developed the complex memorisation strategies required.

However, there are several cognitive skills beyond those described above which influence children's ability to understand and apply ground rules which typically develop during childhood, including theory of mind (ToM) – a concept which describes an individual's ability to understand and interpret the mental state of others (Premack & Woodruff, 1978). Without the ability to understand the mental state of the interviewer, a child is unlikely to comprehend that they are naïve to their experiences, or may hold a false belief about what had happened, which means that they are unlikely to benefit from being instructed to use the "that's wrong" ground rule (e.g., see Koenig & Harris, 2005; Waterman et al., 2004; Welch-Ross, 1999). These skills typically develop from around 4-6 years of age (Peterson et al., 2012; Wellman & Liu, 2004) and should be considered when teaching children to use ground rules.

When considering children's ability to understand and apply ground rules, metacognition should also be addressed. Metacognition is the ability to reflect on your own thoughts and internal states (Flavell, 1979). In order for children to comprehend whether they possess the required information to answer a question asked of them, or if they understand the

question, they need to have mastered these metacognitive abilities including metamemory (reasoning about memories) which involves understanding how to access one's own memories, understanding memory, and how this process can be influenced (Flavell & Wellman, 1977). Fritz et al. (2010) provide evidence that metacognition development increases rapidly during the elementary school years, finding a marked increase in metacognitive performance in 6-8-year olds; and other research has demonstrated that metacognition is not fully developed until around 12-13 years. For example, Markman (1979) found that when children aged 11 years listened to stories that were missing critical information required for adequate comprehension, they were able to demonstrate excellent recall for details of the stories but most failed to identify that their comprehension of the story was inaccurate. Moreover, it was found that children could not correctly identify why a false report was given by a protagonist following a suggestive interview (London et al., 2011). These studies suggest that even when children who do not yet possess these metacognitive abilities are explicitly given a ground rule, they will not be able to use them accurately until these abilities are fully developed.

Once these cognitive abilities are developed, the efficacy of ground rule instruction rests on the development of other advances. This includes the ability to hold the ground rule in the mind during the interview and to implement it where necessary, while inhibiting the urge to respond in inappropriate ways (e.g. by attempting to answer a question that they do not fully understand). This ability to exercise inhibitory control and effective working memory strategies continue to develop into adolescence (Shing et al., 2010) which suggests that even children who demonstrate understanding of the ground rules will not benefit from ground rule instruction due to lacking the executive skills required to execute them adequately.

Saywitz (2002) discusses the developmental underpinnings of testimony, and how communicative competence develops through childhood. Saywitz outlines how children are not able to communicate their inabilities to answer questions effectively until age 10-12, which further raises concerns about younger children's ability to use ground rules throughout an interview. Developmental psychology research reveals that young children are particularly vulnerable to suggestibility due to difficulties with free recall. This makes it tricky to question children without using specific cued questions, increasing the likelihood that adults will use a biased/leading questioning style. Children are also susceptible to suggestibility because they are not apt in accurately identifying the sources of their information (Saywitz & Lyon, 2002). This highlights the importance of investigating best ways to teach children to use the "that's wrong" ground rule, so that children understand that they are encouraged to tell the interviewer if they have made a mistake during an interview. Hence, it is important to discuss whether we can establish a method to teach children to use ground rules that will facilitate children's ability to use them fluently throughout an interview.

How Ground Rules are currently Taught to Children

At present, ground rules are taught to children via verbal instruction, with a brief practice phase during the rapport-building portion of the interview. This involves testing children's understanding of what it means to tell the truth, and teaching them to use the three rules ("I don't know", "I don't understand", "that's wrong") then offering up two example questions for each rule (e.g. "If I said what is my dog's name, what would you say?"). Given the known developmental changes and limitations discussed above, it is important to address how children can learn and apply these ground rules within a forensic interview.

Why Visual Aids could help Children to use Ground Rules

As outlined above, children do not fully develop their inhibitory control and working memory abilities until adolescence which means that children lack the necessary skills to

hold the ground rules in their mind and apply them as necessary, and to supress their urges to attempt to answer questions even when they don't understand the question or know the answer. Perhaps including a visual reminder of the ground rules which will remain accessible to children for the duration of the interview will help remind them of the rules and their purpose and provide support to use them appropriately throughout the interview. Theoretically these visual aids may provide relief from the heavy cognitive load required to retrieve and implement ground rules as there is less need for them to memorise the rules after they have learnt them, and thus less pressure on their working memory capacities.

While there is no known psychological research looking at the use of visual aids as a means of teaching children to use ground rules, there is some evidence to suggest that younger children could benefit from the addition of visual aids as they may help remind the child of the rules and how they are supposed to use them. Visual aids may also provide a non-verbal communication option for those who are less confident about not responding because they have been socialised to believe that adults expect them to communicate in this manner.

The use of visual aids as a teaching tool has been well explored in educational psychology. Shabiralyani et al. (2015) revealed that both teachers and students alike have positive perceptions of using visual aids (such as pictures and animated videos) as teaching tools in the classroom, and that when implemented effectively, visual aids stimulate student thinking and improve learning.

Communication assistants frequently utilise visual resources to assist vulnerable people interacting with the justice system. For example, Legal Aid Western Australia (2016) produced the 'Blurred Borders' visual resources which include storytelling and visual art to help explain the legal concepts around family violence, child protection, bail, the criminal process, etc. Similarly, the Triangle group developed an image vocabulary for children to help them to communicate about sexuality, personal care, feelings, and rights and safety

(2012) which are broadly used in the United Kingdom in several settings. This same group also produced a set of three 'rule cards' designed as reminders to say if you don't know, say if someone gets it wrong, and no guessing (see Collins et al., 2015). However, the efficacy of these rule cards, and the frequency of their use is unclear from the literature which reinforces the need for further research in this area.

The Narrative Elaboration Technique is an investigative interviewing procedure which provides reminders of certain categories about an event that the interviewer would like the child to report on (participants, setting, actions, conversation, and consequences) in the form of line drawings on a card (e.g. see Brown & Pipe, 2003; Saywitz & Snider, 1996). These studies demonstrated that when children were questioned about an event they had participated in, they were able to recall more details about the event during training when the instruction was presented alongside visual cue cards. However, the effects of this training did not carry over into the main memory interview and only seemed to be 'activated' when the visual reminders were present and available to them. This research highlights the importance of the salience of visual aid resources regarding their effectiveness in a forensic interview.

This study proposes that children's inability to retain ground rule instructions and draw on them when they are later required to use them could explain children's consistent failures to accurately use ground rules in an interview. Visual cue cards may therefore ease some of the working memory load of retaining the ground rules whilst doing another task. Visual cue cards may also remind children of the different expectations and their permission to respond to questions in a way not usually accepted by adults.

Paivio (1986) developed the dual coding theory, essentially stating that the human brain utilises different cognitive mechanisms to process visual and verbal information. Paivio's dual coding theory asserts that using both visual and verbal aids simultaneously can have a positive effect on learning using these different mechanisms. A body of research has

been developed that demonstrates the validity of this assertion (Mayer & Sims, 1994), namely that learning is enhanced with both visual and verbal aids. This would then suggest that visual cue cards would have a positive effect on children when learning ground rules because they would provide two potential sources for retrieving the information required to utilise the ground rules.

Dual coding theory suggests that learning can be enhanced by forming mental images because verbal and visual information are processed differently; and can therefore both be used to represent information simultaneously. When retrieving memories, information encoded both visually and/or verbally can be used (Sternberg & Sternberg, 2016). For example, if a child is taught to say "I don't know" when they are unsure of an answer to a question, they may encode the verbal instruction. If they are presented with a visual representation of this same instruction, then they may encode it in both visual and verbal formats, providing multiple retrieval pathways. When it comes to retrieving this memory, a child who was presented with both verbal and visual stimuli during learning has more opportunities for accurate recall, because they have dual codes available to them to access the information.

The use of visual aids is not novel in the realm of forensic interviewing. Wolfman et al. (2018) examined the use of visual aids in interviews (including memory aids such as dolls, drawings, and props) finding that 62% of all interviews used at least one form of visual aid. However, these visual aids were not seen to improve children's performance, but rather they appeared to hinder it as they distracted or confused children from the task at hand. Hence Wolfman et al. suggested that the use of visual aids be kept to a minimum and used with caution. Other research has looked at the use of visual cue cards. For example, Brown and Pipe (2003) investigated children's elaborative reports, and the effectiveness of visual cue cards presented as prompts or reminders of the event that the children were trying to recall.

This study revealed that children responded similarly to either verbal or visual cues, suggesting that it was the opportunity to respond to categorical cues which enhanced their recall, rather than the implementation of visual reminders.

It has been considered whether there may be some risk associated with the use of visual cue cards, namely that they may lead to children overusing the ground rules because they are readily available to them, and they may take it as an 'easy way out' to avoid answering questions, especially during a lengthy interview about a topic which may not be stimulating for a child. Research has shown similar effects in the past. For example, early research on the use of the "don't know" rule showed that children overused it to the detriment of their accuracy because although the inclusion of the ground rule instruction did increase the number of "don't know" responses, they said "I don't know" to questions that they could have answered (Gee et al., 1999; Saywitz & Moan-Hardie, 1994). Researchers adjusted for this by explicitly instructing children to respond to questions that they understood and did know the answer to (Brubacher et al., 2015). Hence, an important part of the current research is to not only test the efficacy of the visual cue cards, but also to identify any unanticipated risks associated with them so that we may produce a good evidence-based guideline for forensic practitioners about when visual cue cards may be appropriate to use.

The current research examines the use of visual cue cards as a means of teaching children to use ground rules during an interview about a past event. The current study specifically investigates: 1) whether using visual aids as a means of teaching children the ground rules during the ground rule training phase will assist children to use the rules more effectively when practicing; 2) whether having visual reminders of the ground rules present during the memory interview will increase children's accurate responding to challenge questions designed to elicit rule responses; and 3) whether having visual reminders of the

ground rules present throughout the memory interview will elicit more spontaneous use of the ground rules during the interview.

The primary aim of this research is to examine how children's ability to use ground rules varies when given normal "business as usual" (BAU) ground rule training compared with using visual aids as a teaching tool, as well as having a visual reminder of the rules throughout the interview post-training. This will be assessed by comparing the performance in responding to scripted 'challenge questions' of children in the BAU condition to that of children in the "visual aids" (VA) condition.

A secondary aim of this research is to determine whether the inclusion of the visual aids helps children to understand and apply the ground rules during the training phase of the interview. This will be examined by comparing children's performance on practice questions delivered to children in the training phase of the BAU condition with performance on the practice questions delivered during training in the VA condition.

A third aim of this research is to examine whether having the visual aid reminders of the ground rules available to children post-training will increase their likeliness of spontaneously using ground rules throughout the interview. This will be measured by comparing the number of spontaneous ground rules used throughout the interview in the BAU condition with the number of spontaneous ground rules uses in the VA condition. The findings of this research will inform us of the best ways to train children to use ground rules to facilitate meaningful conversations and interactions between children and adults.

Children aged 5-12 years participated in a staged event and were then interviewed about it in one of two conditions 1) "visual aids" (VA) and 2) "business as usual" (BAU). These interviews consisted of a ground rule training phase where children were taught the ground rules either verbally (BAU) or verbally + visual aids (VA). Children's use of rules was tested by coding their responses to a set of challenge questions that were designed to

elicit a particular rule response, as well as by noting any spontaneous use of the rules occurring throughout the broader interview.

This study makes a number of predictions, namely that based on dual coding theory (Paivio, 1986), we predict that children instructed with visual cue cards will more effectively learn and utilise the ground rules, demonstrated by higher scores on each of the practice questions, relative to the BAU group. This is because children in the VA group will have access to two sources rather than one when retrieving the ground rule instructions, and because they will have a salient visual reminder of the rules accessible to them throughout the interview which children in BAU do not. This is supported by existing literature that has found visual aids have produced positive outcomes in enabling children to learn and utilise new knowledge (e.g. Dolati & Richards, 2011; Macklin, 1996; Sternberg & Sternberg, 2016). More specifically, this study predicts: 1) that children's performance in responding to practice questions during training will be elevated in the VA condition compared to the BAU condition; 2) that children in the VA condition will perform better in their responses to scripted challenge questions than those in the BAU condition; and 3) that children in the VA condition will be more likely to use ground rules spontaneously throughout the interview than those in the BAU condition because they will be constantly reminded of the rules throughout the interview, and they have a non-verbal option available to them to respond to questions which may support children to elicit more spontaneous uses of the ground rules who may otherwise struggle to let an interviewer know that they do not understand, don't know the answer, or that they have said something incorrect.

Study 1: Pilot Study

The overall aim of this research was to evaluate whether picture cue cards support children's learning and use of ground rules. These cards may help children to learn the rules during the preparation phase of the interview and then serve as a visual reminder or memory

cue for the instructions during the interview, thereby increasing children's spontaneous use of them, and accurate use of them when asked challenging questions designed to elicit particular ground rules responses. Before testing the impact of picture cue cards on ground rule instruction, we wanted to establish children's preference for how the rules ought to be represented. By involving children in the development of the study materials, we hoped to increase the likelihood that the visual aids chosen would be meaningful and effective for them.

Hyde et al. (2014) investigated how children would interact with a realistic or animated person. Children engaged in two conversations with an adult – once where she appeared as herself on video and once where she appeared as a cartoon character. Results indicated that children's interactional style between the two types of conversations did not differ, even for those children who preferred one type of video over the other, suggesting that children do not have an implicit preference for animated versus realistic video imaging.

However, some research suggests that children may prefer realistic photographs over illustrations. Rudisill (1952) had children and adults indicate their preference between realistic drawings and outline drawings, and between realistic photographs and outline drawings that may be used in children's illustrated books. The realistic drawings and the realistic photographs were chosen in preference to the outline drawings, regardless of age.

Given that there has been no prior research in the context of forensic interviewing, this is an exploratory pilot study which assesses children's responses and preferences for picture type and will aid in the development of the study materials for our experimental research to follow. We trialled the use of realistic photographs and emoticons of children depicting the gestures associated with each of the three ground rules. Children were required to choose one or the other when answering tricky questions to determine which type of image children prefer to use.

Method

Design

The current study used a within-subjects design, with one experimental condition. All participants were asked all questions, and all participants saw all images.

Participants

Ethical approval was granted under delegated authority from Victoria University of Wellington's School of Psychology Human Ethics Committee (approval number 00000026259). To compute the minimum number of participants required for this study, an a priori G*Power analysis was used (Faul et al., 2007). This suggested that a sample of 54 participants would be required to detect medium effects (f = .25) with 95% power using a repeated measures ANOVA with alpha at .05.

The lead researcher contacted schools in the wider Wellington region by email and/or phone. After correspondence and meetings with a couple of the schools of interest, it was decided that only the one preschool and primary school would be required to gain a large enough sample size given the power analysis outputs.

A total of 191 participants aged 4-12 years old were recruited from one preschool and one decile 4 primary school in Kāpiti, a suburb of the wider Wellington region in New Zealand. 5 participants were excluded because their responses were unclear, leaving a total sample size of $N = 186$ children. Given this large sample size, we are confident that we had sufficient power to detect any reliable differences in children's responding as a function of image type. The school had a culturally diverse roll, with approximately one third of the students identifying as New Zealand's indigenous culture (Māori). Demographic information about the children at the preschool was not made available to the research team.

Consent was gained from the School Principal, and the Preschool Manager, and letters outlining the study were sent home to the caregivers of all target children explaining how the

study will run. Opt-out consent forms were sent out with these letters, and caregivers were asked to return these within one week if they did not wish for their child to participate in the study. There were no parameters for stopping data collection because the study was delivered to an entire classroom at a time. Therefore, data was collected from all consenting children in the primary school, and for every consenting child in the 4-year-old's classroom at the preschool.

Only two opt-out consent forms were received for two children at the Primary school, so all children who were present in each classroom on the day of the study participated except for the two non-consenting children who completed an independent school task at the side of the classroom.

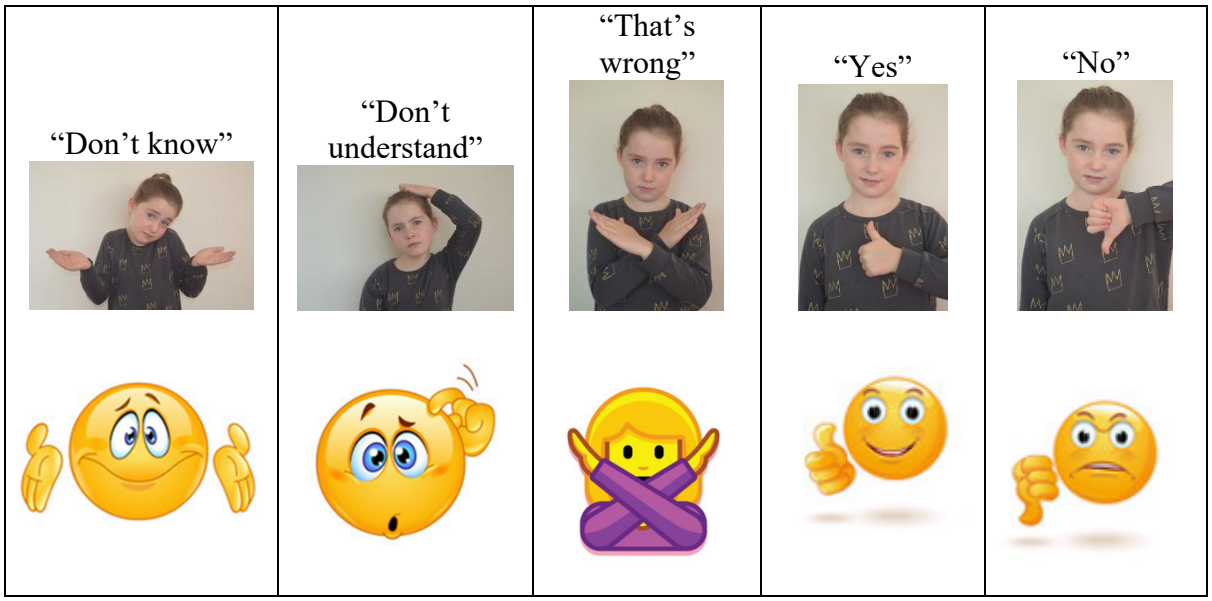
Each child was offered a choice of a small gift (e.g. a pencil, rubber, or stickers) as a thank you for their participation, and each participating school was given a voucher for a local retailer. The cost of the voucher was worked out to be at approximately $10 per participating classroom with a minimum of $20 for any one school.

Materials and Procedure

The lead researcher visited the school and preschool and delivered the questions to the children in their classrooms. No individual time was spent with any child. At the beginning, the researcher gained verbal assent from the participants and handed out a booklet of response sheets to each participant. This booklet contained a series of pictures for each of the 11 questions that would be asked of the children. The children were asked to have a look at the pictures and were given a minute to think about what they meant. The pictures chosen were realistic photographs of a child, and emoticons depicting gestures which reflected the 3 ground rules of interest. For example, a girl shrugging her shoulders to mean "I don't know", scratching her head to mean "I don't understand" and crossing her arms like a "X" in front of her body to indicate "that's wrong" See Figure 1. for the pictures used.

Figure 1

Photographs and emoticons used on response sheets



The researcher explained to the children that she would be asking them some questions, but they did not have to answer verbally, they had to circle one image on each piece of paper that would best help them to answer that question. Of the 11 questions asked of the children, 2 were simple "Yes/No" type questions to determine if the children were able to engage in the task. The remaining 9 comprised of questions assessing the 3 ground rules of interest (3 "don't know" (DK) type questions, 3 "don't understand" (DU) type questions, and 3 "that's wrong" (TW) type questions). For a list of the 11 questions asked of the children, see Table 1. At the end of the session, the researcher asked the children to write their age (without any identifying information) on the front of their response sheets and collected them. Only the lead researcher was involved in data collection, coding, and data entry.

Table 1

Types of questions asked

Question	Question Type
What is my dog's name?	DK

What do I have in my pocket?	DK
What did your teacher have for breakfast this morning?	DK
How old is the scatterbox?	DU
Where in your bedroom?	DU
Why is the scatterbox happy?	DU
What did you get for your 2nd birthday, which was yesterday?	TW
What did you do when the Prime Minister came to school last week?	TW
Did you enjoy your flight home from America this morning?	TW

Preference for image type was determined based on participants' responses. As there was a choice between a realistic photograph and an emoticon response available for each question, preference was determined by which image type the child chose to circle to answer each question respectively.

Correctness of responding was also determined based on participants' responses. For each question asked, the child could choose from a series of images of gestures which either corresponded with the correct response that they should give based on the question type or did not correspond. For example, when asked a TW type question, in order to answer correctly, children should select an image that represents a "that's wrong" gesture (crossing arms in an "X" shape). If they selected either of the "that's wrong" images, then their response was deemed correct. However, if they selected any of the images that corresponded to the other question types (DK or DU gestures) then their answer was deemed incorrect.

Results

Initially, we looked at whether or not children in this study demonstrated a preference in their responding for realistic photographs or emoticons. On each trial, an emoticon choice was given a score of zero, and a photo choice was given a score of one. Scores across the

nine trials (3 trials for each ground rule) were averaged for each participant to produce a possible range of 0 - 1 with scores less than .5 indicating a preference for emoticons, scores greater than .5 indicating a preference for photographs, and .5 indicating no preference.

The mean preference was .24 across the sample (SD = .30) indicating a preference for emoticons. To see whether preference was predicted by age, we conducted a linear regression on the average preference scores across the trials. The model was not significant: just 1.9% of the variance in children's preference for picture type was explained by age (R^2 = .019, $F(1, 184) = 3.50$, $p = .063$).

Next, we considered how correct children's responses were. Across the nine trials children were given a score of 0 when they chose an image that correctly corresponded with the question type, and a score of 1 if they chose an alternate response. Children's responses were averaged across the nine trials, and scores were expressed as a proportion of trials correctly answered, producing a possible range of 0 - 1 with a score of 0 indicating perfect correctness across all trials.

The mean correctness score was .46 (SD = .18) indicating that children chose the correct ground rule image on approximately half of the trials. We examined whether age predicted the accuracy of children's responses with a linear regression. The model was significant. Accuracy increased with age R^2 = .197, $F(1, 186) = 45.63$, $p < .001$, with 19.7% of the variance in accuracy explained by children's age. For each increase in year of age, children's correctness in responding to the questions increased by 3%.

We also considered whether there were differences in children's correctness when responding to the different types of questions. For DK questions, children had a correctness score of M = .37, SD = .35. For DU questions, children had a correctness score of M = .60, SD = .32, and for TW questions children had a correctness score of M = .59, SD = .34. To determine if these means differed significantly from each other, we ran a series of paired

samples t-tests which revealed that the correctness scores for the DK type questions were significantly different from the DU type questions ($t(187) = -6.79$, $p < .001$), and the TW type questions ($t(187) = -6.24$, $p < .001$). Whereas the correctness score for the DU type questions was not significantly different from the correctness score for the TW questions ($t(187) = .27$, $p = .79$).

We used children's averaged correctness scores from the three trials for each rule type in a repeated measures ANCOVA with rule type (DK, DU, TW) as the within subjects factor and age as a covariate. The ANCOVA indicated no main effect of rule type $F(2, 184) = .30$, however a significant rule type by age interaction was observed $F(2, 194) = 3.68$, $p = .027$ partial eta squared = .038. A significant main effect of age was observed, replicating the results of the regression analysis $F(1, 185) = 32.71$, $p < .0001$ partial eta squared = .15. The rule type by age interaction was explained by a significant, positive effect of age on children's correctness when responding to the DK and TW questions (children got more accurate as they got older), but no difference for the DU questions.

Lastly, we considered whether one of the rule type responses was used more frequently than others or if use was similar across the 3 rule types. Across all trials and participants, the DK response was used 758 times, the DU response was used 513 times, and the TW response was used 395 times, despite equal numbers of "correct" opportunities given to use each of the rules. This highlights a potential overuse of the DK rule type.

Overall, the yes/no type and the DK type questions were much more likely to be answered correctly, while the DU and the TW questions were much more likely to be answered incorrectly.

Discussion

This pilot research aimed to determine which type of visual aids would be most appropriate to use with children during a forensic interview. We trialled the use of realistic

photographs and cartoon-like emoticons with children aged 4-12 years while answering tricky questions designed to elicit a specific rule response ("don't know", "don't understand", "that's wrong").

Children's response choices showed an overwhelming preference for emoticon images over realistic photographs, regardless of age. We also found that children were remarkably poor at spontaneously selecting the "correct" image type to use without given any instruction on what the images meant or represent but did observe that children's correctness increased with age. Children were seen to select the appropriate image type more often when responding to DK questions than DU and TW questions, and displayed a tendency to overuse the DK response in general.

When considering the correctness of children's responses, we saw that children were not very accurate in their responding. Children were slightly more likely to correctly respond to the DK type questions than the DU and TW type questions, and while older children were more likely to be correct than younger children, the effect was quite small. For example, when asked "What did your teacher have for breakfast this morning?" they should have selected an "I don't know" picture to use to help them answer that question, or when they were asked a question that they did not understand, e.g. "How old is the scatterbox?" they should have selected an "I don't understand" picture to help them answer. However, this was not the case. The results suggest that children were selecting images almost at random, with a tendency to use the "I don't know" type pictures more often, which has been seen in previous research. For example, McWilliams et al. (2021) investigated the DK response to directive questions in maltreated children following a variety of ground rule instructions. This research found that although the DK instruction reduced children's inaccurate responses in appropriate contexts, it also produced less accurate responding because children were more likely to overuse the DK rule when not appropriate.

Similarly, Henderson and Lyon (2021) investigated whether including a practice question alongside the "I don't understand" ground rule instruction would support children to accurately use the DU response more often. This research found that while older children gave the correct DU more often after being instructed with a practice question, younger children were more likely to give a "don't know" response, again signalling a tendency for children (especially younger) to overuse the DK rule. These findings are congruent with Dickinson et al. (2015) who found that while few children answered all the questions asked of them correctly, some ground rule questions were easier to comprehend than others. Their findings suggested that the DK ground rule and the "tell the truth" ground rule (not assessed in the current study) were the easiest for children respond to accurately.

The relatively low use of the TW response in the current pilot research also suggests that children may be experiencing some difficulty with this rule, or the corresponding image esture chosen to represent the rule in this context. The research team collectively decided on 3 gestures which were thought to map on to each of the ground rules that are included in forensic interviewing protocols - shrugging shoulders to mean "I don't know", scratching head to mean "I don't understand" and crossing arms to indicate "that's wrong". Children's poor accuracy performance may have reflected some ambiguity in the gestures we selected to represent the ground rules. For example, shrugging shoulders for "I don't know" could potentially also be used to say "I don't understand". Alternatively, perhaps children did not recognise the communicative intent of the gestures. As a research team, we agreed that arms crossed in an "X" was a representative gesture of saying "that's wrong", but we cannot be sure that children also thought that this is what the crossed arms gesture meant. We might have eliminated this counter-explanation by asking children to write what the character in each picture was saying, to check that their interpretation of each gesture was similar to ours.

These results suggest that the majority of children do not spontaneously use the "correct" image that corresponds to the ground rule without any context given, which provides further support for the need to educate and train children on how to use ground rules at the beginning of an interview.

In our next study, when using visual cue cards to teach ground rules, we will be providing explicit instruction to children on what each of the gestures mean. This study will be guided by the results of this pilot research, informing us that the best format to design our visual cue cards would be to use cartoon-like emoticons, as we found overwhelming support for children's preference for emoticons over realistic photographs when providing a nonverbal response to difficult questions, and saw (as in previous work) that children tend to struggle with answering difficult or unanswerable questions, in the absence of instructions or training in how to manage them. Our study will therefore employ emoticons displaying gestures of the rules to test whether they support children to learn and employ ground rules when asked difficult questions about a personally experienced event.

Study 2 Method

This Master's research comprises part of a broader research project investigating children's understanding and use of ground rules in investigative interviewing (the Ground Rules and Children's interviews (GRACI) project). Ethical approval was granted under delegated authority from Victoria University of Wellington's School of Psychology Human Ethics Committee (approval number 00000026259). My contribution to the GRACI project examines the use of visual aids as a means of teaching children to use ground rules. I contributed to staging events in schools, conducting memory interviews with children, transcribing interviews, developing a coding scheme, and establishing coding reliability with another independent coder, and coding the data.

Design

A mixed methods design was adopted to compare whether children in the visual aids condition (VA) utilise ground rules more effectively during the training phase and the memory interview than those in the business as usual condition (BAU). Analyses were conducted with the interview condition as a between-subjects independent variable, and rule type as a within-subjects variable. Responses to ground rule training questions during the training phase of the interview, and to challenge questions during the memory interview phase, were used to evaluate the effectiveness of the visual aids in improving children's performance with ground rule responses relative to those in the BAU condition.

A mixed methods design was also used to compare children's spontaneous use of ground rules throughout the interview, and to determine if children in the VA condition were more likely to use ground rules spontaneously than those in the BAU condition; with the interview condition as a between-subjects independent variable, and rule type as a within-subjects independent variable. All spontaneous use of ground rules throughout the interview that were not in response to training questions or challenge questions were tallied according to ground rule type and recorded.

Participants

To compute the minimum number of participants required for this study, an a priori G*Power analysis was used (Faul et al., 2007). This suggested that a sample of $n = 73$ participants would be required to detect large effects ($f = .40$) and a sample of $n = 179$ participants would be required to detect medium effects ($f = .25$) with 95% power using an ANOVA with alpha at .05. Due to disruptions in schools brought on by the global Covid-19 pandemic, data collection was disrupted substantially. Hence, the overall sample size was lower than planned. 89 children (n(male) = 38, n(female) = 50) between ages 5-12 years (M = 8.91 years, SD = 2.02, refer to Table 2) were recruited from primary and intermediate schools across the greater Wellington region to participate. Children were quasi-randomly

assigned to one of two interview conditions, controlling for gender and age, resulting in 44 children in the BAU condition, and 45 in the VA condition. Written parental consent was obtained for each child for both the staged health and safety event, and for the memory interview. Child assent was obtained for the memory interview. Each school received gift vouchers and a morning tea for the staff to thank them for their participation. Each participating child received a novelty gift at the end of the staged health and safety event, and a different gift at the end of the memory interview. Children were offered a choice of a small stationary item such as an eraser or a pencil.

Table 2

Age and sex breakdown

Age (years)	Female	Male	Total
5	6	4	10
6	5	5	10
7	6	4	10
8	6	9	15
9	10	6	16
10	10	2	12
11	6	4	10
12	1	5	6
Total:	50	39	89

Table 3

Age and condition breakdown

Age (years)	BAU (Control)	VA (Visual Aids)	Total
5	5	5	10
6	6	4	10
7	5	5	10

8	8	7	15
9	7	9	16
10	6	6	12
11	5	5	10
12	2	4	6
Total:	44	45	89

Procedure

Health and Safety Event

A Health and Safety demonstration was carried out in each of the recruited schools. These events were run by postgraduate students working in Dr Deidre Brown's Applied Developmental Laboratory and volunteer research assistants during school hours. Participants were randomly assigned to teams of 5-8 and participated in two of four activity stations (Care of cuts, Temperatures, Heartbeats, Hazards) to keep time out of class to a minimum, and to provide 2 non-experienced events to be the focus of challenge questions. At the Care of cuts station, children watched an animated clip about someone getting injured, and were then taught how to treat a minor wound. At the Temperatures station, children were shown how to check their temperature using thermometers. Children partnered with another child and practiced checking each other's temperature under the arm, leg, and in the ears. At the Heartbeats station, children were taught how to check their heartbeat and find their pulse. Children partnered with another child and practiced on each other's chest and stomach, and practiced finding a pulse on their partner's wrist and ankle. At the Hazards station, children were shown pictures of hazardous scenes, identified what the hazard was, and how they could make it safer.

Towards the end of the second activity station, an interruption was staged by an unfamiliar research assistant, who approached the main event leader and engaged in a brief

argument about using the health and safety equipment. This argument was resolved by each of the station leaders providing spare equipment for the interrupter to take with them and the health and safety event continued. Each event lasted approximately 40 minutes.

Follow-up Memory Interview:

Follow-up memory interviews were conducted 14 to 26 days (M = 17.07, SD = 2.92) after the Health and Safety event, depending on child and interviewer availability. In the BAU condition, the delay ranged from 14 to 26 days (M = 17.25, SD = 2.93). In the VA condition, the delay ranged from 14 to 26 days (M = 17.02, SD = 2.77). The delay between event and interview was similar across the two conditions with no significant difference in delay observed $t(86) = .36$, $p = .72$.

Each participant was interviewed individually by one of two trained research assistants, in a private room at the child's school. All interviews were video recorded so they could later be transcribed and coded. Interview duration across interview conditions ranged from 20 minutes to 78 minutes (M = 33.98, SD = 8.61). In the BAU condition, interview duration ranged from 23 to 78 minutes (M = 34.73, SD = 8.36). In the VA condition, interview duration ranged from 20 minutes to 51 minutes (M = 33.23, SD = 7.23). Interview duration was determined by the talkativeness of the child and was similar across the two conditions.

The interviewer collected each child from their classroom and took them to the interviewing space. The interviewer explained that the conversation would be recorded, and individual assent was obtained from each child. Children aged 5-8 provided verbal assent and those aged 9-12 provided written assent. Each interview followed the same structure, with the only difference being the type of ground rule training that each child received. Those in the BAU condition received verbal instruction of the ground rules, while those in the VA condition received verbal instruction accompanied by the visual cue cards.

Ground rule training phase:

The three rules addressed in the current study were to (a) say 'I don't understand', (b) say 'I don't know', and (c) say 'that's wrong' as needed (see Appendix A). Children were also instructed to only tell the truth; however, this rule was practiced but not analysed as part of this research.

Each interview began with ground rule training. The interviewer introduced the child to each of the ground rules in the same order each time, and provided a brief explanation for why/when they should use the rule (e.g., in the BAU condition "…if I ask a question that you don't understand or you're not sure what I mean, just say 'I don't understand', okay? [pause]. If I don't understand something that you say, then I'll let you know, okay? So, let's practice…"); (e.g., in the VA condition "…if I ask a question that you don't understand or you're not sure what I mean, just say 'I don't understand', or you can point to this card here [places 'I don't understand' card on table]. This card helps us when we want to say 'I don't understand', okay?")

Children were given the opportunity to practice using each rule two times during the ground rule training phase, with feedback and reinforcement as necessary. If the child failed to use the rule accurately after these practices, the rule was repeated verbally, and the interviewer moved on to the next rule, or to the beginning of the practice narrative phase after the final rule had been taught and practiced.

Practice narrative phase:

Following the ground rule training phase, the interviewer practiced talking about a recent past event with the child, to assist in building rapport, encouraging elaborative recall, and practicing the open-ended style of interviewing used in the memory interview phase as modelled in Brown et al. (2013). The interviewer asked children to tell them about their

morning and used children's statements as the basis for further open-ended prompting to encourage a detailed account.

Memory interview phase:

The interview began with the prompt, "I heard that a couple of weeks ago, some people came to your school and you did some health and safety activities. Tell me everything you can remember about when that happened". Children were encouraged to elaborate on their own statements using cued invitations (e.g., "You mentioned that you were put into teams, tell me more about when that happened"). For more specific clarifications, closed and directive questioning was used (e.g., "You mentioned that you used different thermometers. How many thermometers were there?"). Responses to closed and directive questions were always followed up with an opened-ended prompt (e.g., "Tell me some more about that").

Challenge questions:

Challenge questions were embedded throughout the memory interview phase, which were designed to test the child's use of the ground rules when they encountered those questions alongside otherwise appropriate questions about a past event. These included questions to elicit a 'don't know' response, e.g., *"You know the girl in the slideshow, was her name Sarah?"*; questions to elicit a 'don't understand' response, e.g., *"When did the bellicose woman leave the hall?"*; and questions to elicit a 'that's wrong' response, e.g., *"Earlier you mentioned that you did different activities, tell me all about what you did at the [WRONG STATION]".* (see Appendix A for further examples of challenge questions asked).

Each child received a total of nine challenge questions in each interview, with three of each ground rule type. Within each rule type, one question was an open-ended prompt, one was a directive (Wh-) question, and one was an option-posing question. The combination of challenge questions received was counterbalanced across conditions. To provide a more naturalistic interview experience the challenge questions were posed at the most appropriate

time point (e.g., when that topic was being discussed), rather than in a single block at the end of the interview as in previous research (e.g., Brown et al., 2019). This style of questioning replicates how a child may be asked questions during a forensic interview, and although this resulted in some variability in the order that some children were asked questions, it provided a more naturalistic and ecologically valid measure of children's responding to challenge questions.

Data Processing

Trained student and volunteer research assistants transcribed all interviews verbatim, recording all verbal and non-verbal responding. All identifying information such as names and locations were removed from the transcripts. Participants' responses to ground rule practice questions during the ground rule training phase, challenge questions during the memory interview phase, and spontaneous use of ground rules throughout the interview were coded according to a coding scheme developed by two student research assistants. Responses to ground rule training questions, and to challenge questions were coded as 'correct response', 'incorrect response', 'abstained response', or 'resisted response' (refer to Table 4). Spontaneous use of ground rules throughout the interview were tallied by rule type and recorded.

Table 4

Response descriptions, codes, scores, and examples

Description	Points awarded	Examples of Participant Responses
Correct response to 'I don't know' challenge question	2	"The woman who came in, did she arrive by bicycle?" C: "Uhh, I don't know"

Correct response to ‘I don’t understand’ challenge question	2	“Did you put the thermometer next to your tympanum?” “I don’t understand”
Correct response to ‘that’s wrong’ challenge question	2	“Where was the man who came in going?” C: “It was a woman”
Resisted response to ‘I don’t know’ challenge question	1	“Tell me what was happening with the leaders before they came and got you” “ummm I don’t understand what you mean”
Resisted response to ‘I don’t understand’ challenge question	1	“When did the bellicose woman leave the classroom?” C: “Mm, I don’t know”
Resisted response to ‘that’s wrong’ challenge question	1	“Was the woman who came in going to Mirimar?” “hmmm I don’t know”
Incorrect response to ‘I don’t know’ challenge question	0	“What was the name of the girl who cut her finger in the slideshow?” “Jessica”
Incorrect response to ‘I don’t understand’ challenge question	0	“Tell me about measuring febrility” “it was really fun”
Incorrect response to ‘that’s wrong’ challenge question	0	“Did you use a red pen to draw a cut on your knee? “Yes”
Abstained response	0	“Tell me about measuring arrythmia”

"After the event we got pencils as a prize"

To test the coding scheme, 15 'training' interview transcripts from the first block of interviewing were used as examples. Within this training phase, disagreements were resolved by discussion and any required rule amendments were made to the coding manual. A further 20 interview transcripts from the first block of interviewing were then independently coded by both coders, and reliability was assessed, using Cohen's Kappa (κ), developed by Cohen (1960) as a measure of the extent to which independent coders assigned the same codes to variables. The κ statistic can be interpreted as having almost perfect agreement when a score between 0.81-1.00 is achieved (McHugh, 2012). Almost perfect agreement was reached between the two coders (κ = 0.93). Further disagreements were discussed, and the coding scheme was amended as needed. The two coders then coded the remainder of the interview transcripts from the first block of interviewing independently, with any unusual responses coded together via collaborative discussion. Of these remaining interview transcripts, 11 were reserved as 'reliability' transcripts to assess the ongoing inter-coder reliability throughout the coding process. They were coded independently by both coders periodically, along with the initial training transcripts. Cohen's κ was calculated regularly as these reliability transcripts were coded, indicating that inter-coder reliability remained consistently high (κ= 0.93) for the first block of interviewing.

After the second block of interviews had been completed, a further 24 interview transcripts were reserved as 'reliability' transcripts. These transcripts were coded independently by both coders, and Cohen's κ was calculated again to determine if inter-coder reliability remained acceptable. Inter-coder reliability remained consistently almost perfect

(κ= 0.96), and the remainder of the interview transcripts were independently coded by the two research assistants.

Scoring responses:

Children's responses were scored according to whether they had responded by using the correct ground rule according to question type (2 points), responding incorrectly by failing to use a ground rule in their response (0 points), abstained from responding (0 points), or resisted by using an alternative ground rule or signalling some trouble with the question (1 point), (see Table 4 for response descriptions, scores and examples).

Results

The dataset consisted of 89 participants with 44 in the "business as usual" control condition (BAU) and 45 in the visual aids condition (VA). Initial inspection of the data indicated that there were missing data points (often due to the interviewer unintentionally missing a challenge question). To determine whether the data were missing at random, Little's Test of Missing Completely at Random (Little's MCAR; Little, 1998) was performed. This MCAR analysis yielded a significant result, indicating that the data was not missing at random. Four of the nine missing data points were coming from one participant because the interview protocol was not correctly adhered to in this case, so this participant was removed from the dataset. A secondary MCAR analysis was run on the remaining missing data points which produced a non-significant result, suggesting that the remaining missing data were missing at random.

To replace the missing values, an Expectation-Maximisation (EM) Imputation was run using SPSS. This is an iterative procedure in which expected values are imputed based on other variables in the dataset, using a value that is the most likely to have occurred naturally (Dempster et al., 1997). This produced a complete dataset with no missing values.

An exploration of the dataset revealed some outliers for the responses to ground rule training questions (n = 13) but because the values of skewness and kurtosis were within acceptable range (Field, 2013), we proceeded with the raw data.

Table 5

Mean accuracy of responses to ground rule training questions, challenge questions, and spontaneous use of ground rules.

Question Type	BAU (Control)	VA (Visual Aids)
Training Questions		
Don't Know (DK)	M = 3.82, SD = .69	M = 3.80, SD = .59
Don't Understand (DU)	M = 3.64, SD = .65	M = 3.43, SD = .93
That's Wrong (TW)	M = 3.93, SD = .33	M = 3.95, SD = .30
Total GR score	M = 3.80, SD = .56	M = 3.73, SD = .43
Spontaneous GR use		
Don't Know (DK)	M = 1.00, SD = 2.31	M = 2.14, SD = 3.30
Don't Understand (DU)	M = .82, SD = 2.76	M = .60, SD = .95
That's Wrong (TW)	M = .82, SD = .95	M = 1.25, SD = 1.16
Total GR score	M = .88, SD = 2.00	M = 1.33, SD = 1.80
Challenge Questions		
Don't Know (DK)	M = 3.23, SD = 1.76	M = 3.59, SD = 1.81
Don't Understand (DU)	M = 4.16, SD = 1.78	M = 4.00, SD = 1.68
That's Wrong (TW)	M = 3.36, SD = 1.48	M = 3.70, SD = 1.21
Total GR score	M = 3.58, SD = 1.67	M = 3.77, SD = 1.57

Children's mean accuracy scores for the questions in the training section of the interview were remarkably low. The maximum score that any child could receive in this

section (regardless of condition) was 18 (there were 9 training questions with each worth up to a maximum of 2 points). Children responded correctly on less than 20% of the trials (M(BAU) = 3.64; M(VA) = 3.66). Inspection of the distributions revealed that the responses to the ground rule training questions were very close to a bimodal distribution where participants tended to either be correct or to make an error in their response. Children who gave any kind of appropriate response to a question (either the correct rule, or an alternate one, or signalling difficulty with the question in some way) were recoded as 1, indicating a correct response, with 0 remaining as an incorrect score.

Chi squares were run on the recoded variables for scores for each ground rule. For the DU ground rule training questions, the chi square was non-significant indicating that the number of correct responses did not differ between the two interview conditions: $X^2(1, N = 88) = .22$, $p = .64$. Participants in BAU obtained a correct score on 73% of the trials and participants in VA obtained a correct score 68% of the trials. For the DK ground rule training questions, the chi square was also non-significant: $X^2(1, N = 88) = .12$, $p = .73$. Participants in BAU obtained a correct score on 91% of the time, and participants in VA obtained a correct score on 89% of the trials. For the TW ground rule training questions, the chi square was again non-significant: $X^2(1, N = 88) = .35$, $p = .56$. In BAU, participants were correct on 95% of the trials, and in VA participants were correct on 98% of the trials. Although the differences between interview conditions were statistically non-significant, looking at percentage performance across the three rule types, children appeared to have more difficulty in correctly answering the DU training questions than the TW and DK training questions.

We examined whether age was associated with performance during training overall and separately for each of the ground rules. Pearson's correlations did not reveal any significant associations between age and performance, either overall or for any of the ground rules considered separately: DU training questions: $r(86) = .08$, $p = .46$; DK training

questions: $r(86) = .01$, $p = .95$; TW training questions: $r(86) = .01$, $p = .96$; overall ground rule training questions: $r(86) = .05$, $p = .62$.

We then looked at spontaneous use of ground rules throughout the interview that were not in response to training or challenge questions. Many upper outliers were noted, with a wide range of values. Skewness and kurtosis were particularly high, so the higher values were log transformed. We then used the transformed variable for spontaneous ground rule use for analysis but report the raw means.

We totalled children's spontaneous use of ground rules (collapsed across rule type) to gain a measure of total overall spontaneous ground rule use. A Pearson's correlation indicated a significant, negative correlation between age and spontaneous use of ground rules $r(87) = -.21$, $p = .046$, indicating as children got older they were giving fewer spontaneous ground rule responses.

To examine whether children differed in their spontaneous use of ground rules as a function of how they had been trained, we conducted a one-way ANOVA with interview condition as the between subjects factor. Spontaneous use of ground rules was relatively infrequent in both conditions with children generally making between 2 and 4 spontaneous ground rule responses across the conditions (see Table 5). The difference between conditions was not significant $F(1, 87) = 2.29$, $p = .13$.

The responses to the challenge questions in the memory interview phase were then analysed. Some outliers were noted ($n = 8$), but because the values of skewness and kurtosis were within acceptable range, and because ANOVAs are robust to violations of normality (Field, 2013), we proceeded with the raw data. Collapsing across rule type, the mean accuracy score for BAU was 3.58 and for VA was 3.77 (see Table 5). Given the maximum score that any child could receive for responses to challenge questions in the memory interview was 18, we can see that in both conditions children's responding was relatively

poor, with few questions responded to with the correct ground rule or any other response signalling difficulty with the questions.

Pearson's correlations revealed that there was no significant association between children's age and their performance on responses to challenge questions, either overall or for any of the individual rules. There was no significant correlation shown between age overall performance on responses to challenge questions $r(86) = -.09$, $p = .40$. There was no significant correlation observed between age and performance on responses to DU challenge questions $r(86) = -.05$, $p = .61$; responses to DK challenge questions $r(86) = -.01$, $p = .90$; and responses to TW challenge questions $r(86) = -.13$, $p = .22$.

To examine whether there were differences according to interview condition and rule type, we performed a mixed measures ANOVA with interview condition as the between subjects factor and rule type as the within subjects factor. There was a significant effect of rule type observed $F(2, 172) = 5.04$, $p = .007$. These results reflect a medium effect size with a partial eta squared of .11.

Post-hoc analysis with a Bonferroni adjustment between responses to different rule types revealed that children's correctness in response to DU challenge questions were not significantly different to their correctness responding to DK challenge questions (0.67 (95% CI, 0.18 to 1.16), $p = .007$) or TW questions (0.55 (95% CI, 0.15 to 0.94), $p = .008$). There was also no significant difference in children's correctness when responding to DK and TW challenge questions (0.13 (95% CI, -0.33 to 0.58), $p = .58$).

There was no significant effect of interview condition $F(1,86) = .61$, $p = .44$ observed, and no interview condition and rule type interaction $F(2,172) = .87$, $p = .42$. Because of the variation in interview duration observed with interviews ranging from 20 minutes to 78 minutes across the two interview conditions ($M = 33.98$, $SD = 8.61$), we ran a t-test to determine if there was a difference in interview length between the two conditions. No

significant difference in interview duration observed $t(86) = .82$, $p = .42$. We also ran a Pearson's correlation to check whether interview duration was related to children's accuracy when responding to the challenge questions, which was also non-significant: $r(86) = .03$, $p = .82$.

These analyses indicate that visual aids did not improve children's performance in responding to difficult questions with a ground rule response, but equally, no negative effects of presenting the visual cues were evident. Rather, all children demonstrated low levels of accuracy across the board when responding to the challenge questions.

Discussion

The purpose of this research was to determine whether visual cue cards are an appropriate tool to teach children to use ground rules, and to support them to use ground rules as needed throughout an interview. The pilot study revealed that children preferred cartoon-like emoticons over realistic photographs, so cue cards were developed using emoticons displaying gestures representing the ground rules. We were interested in whether children in the VA condition (who received visual cue card instruction) would give more correct responses to training questions compared to children in the BAU condition (who only received verbal instruction). We were also interested in whether having cue cards would support children to respond to challenge questions appropriately during an interview about a recent event, and if they would encourage more spontaneous use of ground rules.

The data did not support our prediction that children's performance during training would be elevated in the VA condition compared to the BAU condition. Children's responding to training questions was remarkably poor across both conditions. We also did not find support for our prediction that children in the VA condition would perform better when responding to challenge questions during the interview. Children's performance on challenge questions was relatively poor across conditions, with very few questions answered in an

appropriate manner. A negative correlation between age and spontaneous use of ground rules was observed, indicating that older children spontaneously used ground rules less than younger children. We predicted that children in the VA condition would demonstrate more spontaneous ground rule use than those in the BAU condition, but no significant differences were observed. Spontaneous use of ground rules was also relatively infrequent across conditions.

Visual Cue Cards and Children's Ground Rule Responding

These findings indicate that visual cue cards do not improve children's performance in responding to difficult questions with a ground rule. Although no facilitative effects were observed, no negative effects of using visual cue cards were evident either. Rather, all children demonstrated low levels of accuracy in responding to questions across conditions. One possible explanation for the lack of notable differences across conditions is that children are still developing important cognitive abilities that may underpin their effective use of ground rules (Brubacher & Dickinson, 2015). One example is theory of mind abilities which are crucial for children to understand the interviewer's state of mind (e.g. Koenig & Harris, 2005). Metacognition is also relevant to ground rule use for children to be aware of and coherently articulate their own thought processes (Flavell, 1979). Inhibitory control is another ability that children may not have sufficiently developed yet which is crucial to ground rule use, as it allows them to filter out irrelevant stimuli and to supress unhelpful behavioural responses. Similarly, working memory abilities may still be underdeveloped which allow children to hold and manipulate multiple pieces of information at once (Shing et al., 2010). Such abilities are important for recognising *when* a ground rule ought to be used, and *what to do* to employ one.

Perhaps visual cue cards simply remind children of what the rules are, and therefore may not sufficiently address the broader cognitive and social processes implicated in ground rule use. An alternative training strategy which incorporates more extensive explanation of, and practice with using the rules may be a more effective way to support children in knowing how to manage difficult questions (discussed further below).

Alternatively, it is possible that these visual cue cards do reduce the cognitive load on children significantly by providing a reminder of the ground rules they ought to use and therefore reducing the pressure on their working memory, inhibitory control, and metacognitions, but this is not the active mechanism inhibiting the understanding and correct use of ground rules. For example, children in collectivistic cultures have been shown to be more compliant and attempt to please the interviewer by giving answers in line with the interviewer's expectations (Mehrani, 2011). Similarly, children from collectivistic cultures, where they are expected to be more compliant or are more likely to agree with adults (Mehrani & Peterson, 2016; Mehrani & Peterson 2018), may be less likely to correct adults and implement the "that's wrong" rule, or more willing to attempt to answer adult's questions rather than use the "that's wrong" rule. In this case, the active mechanism inhibiting the correct use of ground rules is not cognitive load, but rather cultural and social factors influencing how children think they should respond to questions from adults. Future research should aim to identify other possible factors that may influence children's ability to use ground rules effectively.

It is also possible that the gestures depicted in the cue cards did not effectively represent the rules for children, and so were not useful in prompting them to use the rules. Some research suggests that gestures are idiosyncratic and are constructed in the moment of speaking rather than reflecting an agreed upon conventional code of language (Goldin-Meadow, 2000). Although the use of cartoon-like emoticons was endorsed by children in the

pilot study, and gesturing is widespread and robust, and occurs across cultures, ages, and tasks (Feyereisn & Lannoy, 1991), there was some ambiguity about the meaning of the gestures. This could be overcome in future by asking children to describe how they would represent each of the ground rules and using this qualitative information to develop alternate cue cards which are representative of a child's view of the ground rules. For example, with digital technology available, children could be asked how they would show that they did not know the answer to a question, and a photograph could be taken (along with pictures of gestures for not understanding or for correcting an error) and displayed on a device for children to point to. Another suggestion for checking children's interpretation of gestures representative of ground rules would be to include it in the training phase of an interview. The interviewer could show a picture card, ask them what it means, then ask them when they could use such a picture or gesture, and then move on to the ground rule instruction and practice phase.

Although there are theoretical explanations for the lack of impact of using visual cues cards on children's use of ground rules, our findings may also reflect a sampling issue. The global COVID-19 pandemic presented the research team with several barriers to collecting data due to nation-wide lockdowns during the data collection phase, resulting in a smaller sample than originally planned for. This smaller sample size could mean that a lack of power precluded detecting effects. This has been noted when interpreting our data, but due to the complete lack of any effects observed, and the uniformly poor performance across the sample, it is likely that even with a larger sample size, differences would not have been detected. As part of the broader research programme, a group of children were recruited and interviewed about the event (with challenge questions) in the absence of any ground rule instruction at all. Analysed of the differences (or lack thereof) between this group and the VA

and BAU conditions would further illuminate whether ground rules provide any support to children at all.

Varying Responses to Different Question Types

Whilst we did not observe differences between interview conditions, there were some indications that children may manage some types of difficult questions more easily than others. The percentage performance scores across the three rule types suggest that children had more difficulty in correctly answering the DU training questions than the TW and DK training questions. This is similar to findings by Dickinson et al. (2015) who's research showed that children's performance during training was superior on the DK questions.

However, during the memory interview phase in our research, children displayed the most accurate responses to DU challenge questions compared to DK and TW. This finding differs to research by Danby et al. (2015) which demonstrated that practice in training of the ground rules only benefitted the use of the DK rule, as children in their study performed better on DK compared to TW and DU. Several reasons for this phenomenon are suggested.

Firstly, we discuss how the rules are in some way inherently different to each other. Brown et al. (2019) looked at the relationships between ground rule performance in children with and without intellectual disabilities, and found that there were no associations between children's responding to different ground rules (DK, TW, DU) during the training phase of the interview. Brown et al. suggested that just because a child has grasped an understanding of one ground rule, does not automatically mean that they will understand the others. This is because children are unable to transfer knowledge between domains (discussed below), which may explain why children perform differently on different ground rules.

We considered whether there is something about the types of questions used during training, which was influencing children's responses, making it more difficult to respond

correctly to the DU questions. An example of a DK training question is “What is my dog’s name?”; an example of a TW training question is “If I said that you are a two-year-old GIRL/BOY [wrong gender], what would you say?”; and an example of a DU training question is “Where is the querulous cat?” (see Appendix B for further examples). On the surface it seems that the DK and TW questions were simpler in the training phase and therefore easier for children to respond to. For example, a child could omit the word “querulous” from the DU question and respond to the amended question “Where is the cat?”, with any answer other than “I don’t understand”, thereby providing an incorrect answer. However, one could argue that it is not as simple to omit information from the TW and DK questions and to answer them outright, providing more opportunities for children to utilise the ground rules correctly in these instances.

Another suggested reason for children’s apparent reluctance to state that they did not understand during training (as indicated by a low correct use of the DU rule in this phase) is that this may be due to children’s socialisation to always answer adults’ questions, even if they are only guessing (Lamb & Brown, 2006). This socialisation may have equally contributed to their likelihood to use the DU rule more accurately than other rules during the memory interview phase because it became more obvious to them that they should state when they do not understand. This is because this is a novel and unusual experience for children.

Transference of Training

Another contributing factor could be the format of the DU challenge questions, which directly mirrored the format of the DU training questions– usually comprising of a single word that is difficult for the child to understand (e.g. querulous in training questions and bellicose in challenge questions; see appendices A & B). As such, the format of these questions may have been easily recognisable to children as something that they should

respond to with a ground rule. This is congruent with research that explains how children learn to transfer training. Day and Goldstone (2012) suggest that surface features (such as the format of a question) help them to recognise when and how to apply their knowledge. Because the format of the DU questions was similar in the training and the interview phases, this suggests that the DU rule may have been more salient to children throughout the interview. This would explain the slightly better performance with the DU rule during the memory interview phase compared to training.

However, the difficulties in the challenge questions for the other rules were more subtle compared to the DU questions. For example, children were asked "Where was the leader wearing her tiger stethoscope?" as a TW question. The challenge to be addressed in this question is that the leader was not wearing a tiger stethoscope, she was wearing a bear stethoscope. Many children answered that she was wearing it "around her neck" which was the correct location of the leader's stethoscope, despite which animal covering the leader was wearing. The inaccuracy in this question is much more nuanced and trickier to detect than a word such as bellicose in the DU challenge questions, so it may not have been as obvious that it required a ground rule response. This suggests that when the surface features required for transference of training (Day & Goldstone, 2012) are not present, children are not applying the ground rules.

Research tells us that children require not only these surface features to adequately transfer their knowledge; they require a combination of surface and similarity features which affords them an understanding of the underlying 'principle' of the task. This allows them to see that they can use these strategies in situations that are dissimilar to the scenarios in which they learnt them (Day & Goldstone, 2012). This is problematic for the field of forensic interviewing because interviewer's questions that require a ground rule response are not intentionally and carefully formulated in a way that is recognisable to children as having

something that is incorrect, incomprehensible, or unknown. More likely, they will reflect subtle nuances and inaccuracies in the interviewer's statements, like the subtle challenges in the TW question example above.

Whilst ground rules *may* be of use, the onus should not ultimately be on children to signal difficulty with interviewer's questioning (this should be the least preferred method of highlighting difficulty). Instead, care and attention should be taken by interviewers to formulate their questions in such a way that children are given the best opportunity to confidently provide meaningful responses. The implication of putting the onus entirely on children to signal difficulty is that if they failed to do so when required during a high stakes interview, communication may go awry, and their evidence and credibility may be undermined. Instead, interviewers should be careful with their questions, probe children's understanding thoroughly, ask for elaboration to be sure the question was understood and answered as intended, and give children reminders and options to decline to provide a response. In other words, a thorough understanding of the correct verbal interview protocol is of utmost importance when interviewing children to gain substantial accurate information about an event (Lamb et al., 2018).

The Validity of Current Ground Rule Training

Children's difficulties in responding to these nuanced type challenge questions suggest that children are not likely to use the rules appropriately when responding to difficult questions if they don't recognise them as something that requires a ground rule response. This raises the question of whether the use of ground rules is redundant, or if the training that we are giving children to use ground rules is insufficient. Perhaps it would be beneficial to incorporate a longer training phase using the questions, and incorporating more subtle and

varied examples of the rules, so that children become familiar with the more nuanced types of mistakes and difficulties they are likely to face during an interview.

Further efforts should be put into helping children to understand the context and culture of the interview space so that they are supported to utilise ground rules effectively. Future research should address this by exploring alternative training and practice methods of using the ground rules in the interview. The broader research project that this master's thesis forms a part of is currently investigating the impact of differing levels of ground rule instruction with children to determine if elaborative training and narrative practice using the ground rules will help or hinder children's ability to use them.

Disparities in Spontaneous Ground Rule Use

Spontaneous ground rule responses were infrequently used by children across conditions in the current research, which suggests that having the cue cards available to them throughout the interview did not encourage children to use ground rules more. It may be that the non-challenge questions and statements were too easy or familiar, and spontaneous ground rules simply weren't required in this context. The children were speaking freely about a salient out-of-the-ordinary event that had occurred only two weeks prior, so maybe they were able to answer questions accurately without requiring assistance from ground rules. Research shows that children can be highly accurate and reliable when recalling past events, especially when interviewed using open questioning (Lamb et al., 2018). In the current research, the interviewers closely followed an evidence-based interview protocol (NICHD; Lamb et al., 2018) which promotes the use of open-ended questions and uses children's responses as cues when probing for further information. This suggests that children in this context were unlikely to encounter statements and questions (outside of the challenge

questions) from the interviewers that required a ground rule response, hence leaving them with fewer opportunities to use ground rules spontaneously (Brown & Lamb, 2015).

Although spontaneous use of ground rules was relatively low across both conditions, older children spontaneously used them less than younger children. Perhaps as children get older, they become more accustomed to answering adults' questions or to filtering out parts of the question that they do not understand or cannot respond to for some reason. Younger children are not likely to have developed this ability to the same extent, and so may be more likely to spontaneously tell the interviewer when they don't know/don't understand. On the other hand, this higher spontaneous use of ground rules among younger children could reflect that younger children were finding the content more difficult and therefore more opportunities to use ground rules were presented to them than older children, despite the interviews being conducted in similar formats.

Is there still a place for Visual Aids in Forensic Interviewing?

Given that no detrimental effects of using visual cue cards were observed in this research, the inclusion of visual aids may still be useful as they may serve other purposes in the interview such as rapport building and communication. For example, research has demonstrated that visual aids have been beneficial for children in other contexts during interviews. Stalker and Connors (2003) found that younger children were able to chat more freely and volunteer more information about difficult subjects such as grieving for a lost grandparent while engaging with visual aids, suggesting that the visual supports provided a comforting environment for children to express their feelings. Stalker and Connors also suggested that the visual supports in their study facilitated communication with children who had disabilities such as a cognitive or hearing impairment. Children with disabilities were not included in the current research, so perhaps a future study could investigate whether our

visual cue cards would help children with cognitive, communicative or physical disabilities to learn and use ground rules.

Research suggests that children with intellectual disabilities (ID) are more likely to be involved with child protection services than children without ID (Slayter & Springer, 2011), yet they face difficulties with communication due to cognitive and social deficits (Brown et al., 2012) and hence they are often viewed as non-credible eye witnesses (Wyman & Talwar, 2019). Perhaps children with ID could benefit from visual supports during an interview to scaffold their own communication competencies and to potentially increase their credibility as an eyewitness. Future research should investigate this with a sample of children with ID.

Summary

We still have much to learn about whether ground rules are an appropriate intervention or protection for children to manage the difficulties they might encounter in an interview and if ground rules are the best way to train children to recognise difficult questions and how to respond to them correctly to ensure that their responses are as complete and accurate as possible. This research demonstrates that current methods of teaching children to use ground rules in a forensic interview are not effective, as evidenced by their poor performance with using them throughout these interviews. We have demonstrated that visual cue cards are not a sufficient support to help children overcome the cognitive load required to learn, retrieve, and apply ground rules effectively. Future research should explore a more comprehensive training method which includes an extended verbal instruction of the ground rules and incorporates more varied and nuanced opportunities for children to practice using the ground rules before the memory interview commences.

References

Aguiar, N., Stoess, C., & Taylor, M. (2012). The Development of Children's Ability to Fill the Gaps in Their Knowledge by Consulting Experts: Children's Ability to Fill the Gaps in Their Knowledge. *Child Development*, *83*(4), 1368–1381. https://doi.org/10.1111/j.1467-8624.2012.01782.x

Brubacher, S. P., Poole, D. A., & Dickenson, J. J. (2015). The use of ground rules in investigative interviews with children: A synthesis and call for research. *Developmental Review*, *36*, 15–33. https://doi.org/10.1016/j.dr.2015.01.001

Brown, & Lamb, M. E. (2015). Can Children Be Useful Witnesses? It Depends How They Are Questioned. *Child Development Perspectives*, *9*(4), 250–255. https://doi.org/10.1111/cdep.12142

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

Brown, D. B., 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

Brown, D., & Pipe, M. E. (2003). Individual differences in children's event memory reports and the narrative elaboration technique. *Journal of Applied Psychology*, *88*(2), 195–206. https://doi.org/10.1037/0021-9010.88.2.195

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Collins, K., Marchant, R., & McAuley, C. (2015). Advocate's gateway toolkit: Planning to question a child or young witness-. https://0abb646f-e7c3-4c9e-b37e-d86ac13cfbb3.filesusr.com/ugd/1074f0_57e3427cfbaa4635abedb100c9d4ae02.pdf

Cooper, A., Wallin, A. R., Quas, J. A., & Lyon, T. D. (2010). Maltreated and nonmaltreated children's knowledge of the juvenile dependency court system. *Child Maltreatment*, *15*(3), 255–260. https://doi.org/10.1177/1077559510364056

Danby, M. C., Brubacher, S. P., 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 & 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 of transfer of learning. *Educational Psychologist*, *47*(3), 153–176. https://doi.org/10.1080/00461520.2012.696438

de Haan, I., Joy, E., Beddoe, L., & Iam, S. (2019). "The tip of the iceberg": Multiple thresholds in schools' detecting and reporting of child abuse and neglect. *Children and Youth Services Review*, *96*, 278–285. https://doi.org/10.1016/j.childyouth.2018.10.034

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B, Methodological*, *39*(1), 1–38. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

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

Dolati, R., & Richards, C. (2011). Harnessing the use of visual learning aids in the English language classroom. *Arab World English Journal*, *2*(1), 3-17.

https://www.researchgate.net/profile/Romana_Dolati/publication/220029726_Harnessing_the_Use_of_Visual_Learning_Aids_in_the_English_Language_Classroom/links/0922b4f44dea9aae4e000000.pdf

Drummey, A., & Newcombe, N. (2002). Developmental changes in source memory. *Developmental Science*, *5*(4), 502–513. https://doi.org/10.1111/1467-7687.00243

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

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioural and biomedical sciences. *Behaviour Research Methods*, *39*, 175–191. https://doi.org/10.3758/BF03193146

Feyereisen, P., & de Lannoy, J. D. (1991). Gestures and speech: Psychological investigations. New York: Cambridge University Press.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London, England: SAGE Publications.

Flavell, J., Friedrichs, A., & Hoyt, J. (1970). Developmental changes in memorization processes. *Cognitive Psychology*, *1*(4), 324–340. https://doi.org/10.1016/0010-0285(70)90019-8

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), 59-68 https://doi.org/10.1177/1077559501006001006

Gee, S., Gregory, M., & Pipe, M. E. (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

Golding, J. M., Lynch, K. R., Wasarhaley, N. E., & Keller, P. S. (2015). Courtroom perceptions of child sexual assault: The impact of an eyewitness. *Criminal Justice and Behavior*, *42*(7), 763–781. https://doi.org/10.1177/0093854814568552

Goldin-Meadow, S. (2000). Beyond words: The importance of gesture to researchers and learners. *Child Development*, *71*(1), 231–239. https://doi.org/10.1111/1467-8624.00138

Ellis, L. M., Powell, M. B., Thomson, D. M., & Jones, C. (2003). Do simple "groundrules" reduce preschoolers' suggestibility about experienced and nonexperienced events?. *Psychiatry, Psychology and Law*, *10*(2), 334-345. https://doi.org/10.1375/132187103322742167

Flavell. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *The American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail Jr. & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3–30). Hillsdale, NJ: Erlbaum.

Fritz, H. P., & Kleitman, S. (2010). "How do I remember when I got my dog?" The structure and development of children's metamemory. *Metacognition and Learning*, *5*(2), 207–228. https://doi.org/10.1007/s11409-010-9058-0

Henderson, & Lyon, T. D. (2021). Children's signaling of incomprehension: The diagnosticity of practice questions during interview instructions. *Child Maltreatment*, *26*(1), 95–104. https://doi.org/10.1177/1077559520971350

Hughes, M., & Grieve, R. (1980). On asking children bizarre questions. *First Language*, *1*(2), 149–160. https://doi.org/10.1177/014272378000100205

Hyde, J., Kiesler, S., Hodgins, J., & Carter, E. (2014). Conversing with children: cartoon and video people elicit similar conversational behaviors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1787–1796. https://doi.org/10.1145/2556288.2557280

Koenig, & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, *76*(6), 1261–1277. https://doi.org/10.1111/j.1467-8624.2005.00849.x

Lamb, M., & Brown, D. (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., Brown, D., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2018). *Tell Me What Happened: Questioning Children about Abuse.* John Wiley & Sons, Incorporated.

La Rooy, D. J., Brubacher, S. P., Aromäki-Stratos, A., Cyr, M., Hershkowitz, I., Korkman, J., Roberts, K. P. (2015). The NICHD protocol: a review of an internationally-used evidence-based tool for training child forensic interviewers. Journal of Criminological Research, Policy and Practice, 1(2), 76–89. https://doi.org/10.1108/JCRPP-01-2015-0001

Legal Aid WA (2016). *Blurred Borders.* Accessed December 8th 2021, Retrieved from https://blurredborders.legalaid.wa.gov.au/

Little, R.J. (1997). Biostatistical analysis with missing data. In: *Encyclopedia of Biostatistics* (ed. P. Armitage and T. Colton). London: Wiley.

London, Bruck, M., Poole, D. A., & Melnyk, L. (2011). The development of metasuggestibility in children. *Applied Cognitive Psychology*, *25*(1), 146–155. https://doi.org/10.1002/acp.1653

Lyon, T. (2014). Interviewing children. *Annual Review of Law and Social Science*, *10*(1), 73–89. https://doi.org/10.1146/annurev-lawsocsci-110413-030913

Macklin, C. M. (1996). Preschoolers' learning of brand names from visual cues. *The Journal of Consumer Research*, *23*(3), 251–261. https://doi.org/10.1086/209481

Mayer, R., & Simes, V. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, *86*(3), 389–401. https://doi.org/10.1037//0022-0663.86.3.389

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

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

Mehrani, M. B. (2011). What is biased? Children's strategies or the structure of yes/no questions? *First Language,31*(2), 214–221. https://doi.org/10.1177/0142723710371886.

Mehrani, M. B., & Peterson, C. (2016). Interviewing preschoolers: Response biases to yes–no questions. *Applied Cognitive Psychology,31*(1), 91–98. https://doi.org/10.1002/acp.3305.

Mehrani, M. B., & Peterson, C. (2018). Responses to interview questions: A cross-linguistic study of acquiescence tendency. *Information Child Development,27*(2). https://doi.org/10.1002/icd.2063.

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

Paivio, A. (1990). Mental representations: A dual coding approach. In *Mental Representations*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195066661.001.0001

Peterson, C., & Grant, M. (2001). Forced-choice: Are forensic interviewers asking the right questions?. *Canadian Journal of Behavioural Science 33*(2), 118 – 127 https://doi.org/10.1037/h0087134

Peterson, C. C., Wellman, H. M., & Slaughter, V. (2012). The mind behind the message: advancing theory-of-mind scales for typically developing children, and those with deafness, autism, or asperger syndrome. *Child Development*, *83*(2), 469–485. https://doi.org/10.1111/j.1467-8624.2011.01728.x

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 4*, 515–526. doi:10.1017/S0140525X00076512

Roberts, K. P., & Lamb, M. E. (1999). Children's responses when interviewers distort details during investigative interviews. *Legal and Criminological Psychology*, *4*(1), 23-31. https://doi.org/10.1348/135532599167752

Rudisill, M. (1952). Children's preferences for color versus other qualities in illustrations. *The Elementary School Journal*, *52*(8), 444–451. https://doi.org/10.1086/459378.

Saywitz, K. J. (2002). Developmental underpinnings of children's testimony. *Children's testimony: A handbook of psychological research and forensic practice*, 3-19.

Saywitz, K. J., & Lyon, T. D. (2002). Coming to grips with children's suggestibility. In M. L. Eisen, J. A. Quas, & G. S. Goodman (Eds.), *Memory and Suggestibility in the Forensic Interview* (pp. 85–113). Lawrence Erlbaum Associates Publishers.

Saywitz, K. J., & Moan-Hardie, S. (1994). Reducing the potential for distortion of childhood memories. *Consciousness and Cognition*, *3*(3-4), 408-425. https://doi.org/10.1006/ccog.1994.1023

Saywitz, & Snyder, L. (1996). Narrative elaboration: Test of a new procedure for interviewing children. *Journal of Consulting and Clinical Psychology*, *64*(6), 1347–1357. https://doi.org/10.1037/0022-006X.64.6.1347

Saywitz, K. J., Snyder, L., & Nathanson, R. (1999). Facilitating the communicative competence of the child witness. *Applied Developmental Science*, *3*(1), 58–68. https://doi.org/10.1207/s1532480xads0301_7

Shabiralyani, G., Hasan, K. S., Hamad, N., & Iqbal, N. (2015). Impact of visual aids in enhancing the learning process case research: District Dera Ghazi Khan. *Journal of Education and Practice*, *6*(19), 226-233. https://files.eric.ed.gov/fulltext/EJ1079541.pdf

Shing, Y. L., Lindenberger, U., Diamond, A., Li, S. C., & Davidson, M. C. (2010). Memory maintenance and inhibitory control differentiate from early childhood to adolescence. *Developmental Neuropsychology*, *35*(6), 679–697. https://doi.org/10.1080/87565641.2010.508546

Slayter, E. & Springer, C. (2011). Child welfare-involved youth with intellectual disabilities: Pathways into and placements in foster care. *Intellectual and Developmental Disabilities*, *49*(1), 1–13. https://doi.org/10.1352/1934-9556-49.1.1

Stalker, K. & Connors, C. (2003). Communicating with disabled children. *Adoption & Fostering*, *27*(1), 26–35. https://doi.org/10.1177/030857590302700105

Sternberg, R. J., & Sternberg, K. (2016). *Cognitive psychology*. Nelson Education.

Waterman, A. H., Blades, M., & Spencer, C. (2004). Indicating when you do not know the answer: The effect of question format and interviewer knowledge on children's "don't know' responses. *British Journal of Developmental Psychology*, *22*(3), 335–348. https://doi.org/10.1348/0261510041552710

Welch-Ross, M. K. (1999). Interviewer knowledge and preschoolers' reasoning about knowledge states moderate suggestibility. *Cognitive Development*, *14*(3), 423–442. https://doi.org/10.1016/S0885-2014(99)00013-1

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Westera, N., Zajac, R., & Brown, D. (2016). Witness interviewing practices in New Zealand. *International developments and practices in investigative interviewing and interrogation*, *1*, 87-98. https://doi.org/10.4324/9781315769684-18

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

Wyman, L. J., & Talwar, V. (2019). Best practices for interviewing children with intellectual disabilities in maltreatment cases. *Exceptionality: the Official Journal of the Division for Research of the Council for Exceptional Children*, *27*(3), 167–184. https://doi.org/10.1080/09362835.2018.1425623

Zajac, R., & Brown, D. (2018). Conducting successful memory interviews with children. *Child & Adolescent Social Work Journal*, *35*(3), 297–308. https://doi.org/10.1007/s10560-017-0527-z

Appendix A

Table of challenge questions used throughout memory interview

Rule type	Question type	Challenge Question
Don't know	Closed	"You know the girl in the slideshow who cut her finger, was her name Sarah?"
Don't know	Closed	"The woman who came in to take the spare equipment, did she arrive by bicycle?"
Don't know	Closed	"Did the girl who ran with scissors fall and cut her brother?"
Don't know	Directive	"What was the name of the girl who cut her finger in the slideshow?"
Don't know	Directive	"What was the leader's temperature?"
Don't know	Directive	"In the picture of the girl with scissors, what happened after she ran with scissors?"
Don't know	Directive	"The woman who came in to take the spare equipment, what colour was her bicycle?"
Don't know	Directive	"How much did the stethoscopes cost the leader at the heartbeat station?"
Don't know	Open-ended	"Tell me everything that happened before the event, when the leaders were getting all the activities set up for you guys?"
Don't understand	Closed	"In the picture of the road, was the man on the zebra crossing running with impigrity?"
Don't understand	Closed	"Did you put your thermometer next to your tympanum?"

Don't understand	Closed	"Did Arthur hurt his patella?"
Don't understand	Closed	"Did the bellicose woman leave the classroom after picking up the spare equipment?"
Don't understand	Closed	"Did someone auscultate your pulse?"
Don't understand	Directive	"When did the bellicose woman leave the classroom?"
Don't understand	Directive	"When did you put the thermometer next to your tympanum?"
Don't understand	Directive	"Which part of his crus did Arthur hurt?"
Don't understand	Directive	"In the picture of the road, why was the man in the zebra crossing running impigriously?"
Don't understand	Open-ended	"Tell me about measuring febrility"
Don't understand	Open-ended	"Tell me about measuring arrhythmia"
That's wrong	Closed	Was your temperature the same *[use higher or low if temp was really the same*] as your partners or [*higher/lower – use incorrect*]?
That's wrong	Closed	"Did you use a red pen to draw a cut on your knee?"
That's wrong	Closed	"Was the woman who came in going to Karori?"
That's wrong	Closed	"Did you like your leader's tiger stethoscope?"

That's wrong	Closed	"In the picture of the garden, did the boy hurt his foot when he stepped on the rake?"
That's wrong	Directive	"Where did you draw the pretend cut?"
That's wrong	Directive	"Where was the man who came in going?"
That's wrong	Directive	"When the leader saw the temperature was 42 (*wrong temperature*), what did she say?"
That's wrong	Directive	"In the picture of the garden, who got hurt stepping on the rake?"
That's wrong	Directive	"Where was the leader wearing her tiger stethoscope?"
That's wrong	Open-ended	"Earlier you said you did different activities, tell me about what you did at the [*wrong station*]"

Each interview contained 9 of the challenge questions listed above including 3 "don't know" type questions which comprised of 1 open-ended question, 1 closed question, and 1 directive question; 3 "don't understand" type questions which comprised of 1 open-ended question, 1 closed questions, and 1 directive question; and 3 "that's wrong" type questions which comprised of 1 open-ended question, 1 closed questions, and 1 directive question.

Appendix B

Table of ground rule training questions

Rule type	Training Question
Don't know	'If I asked you, "What is my dog's name?" What would you say?'
Don't know	'If I asked you "what were you doing on the 1st June [*use another date if this is their birthday*] two years ago?" What would you say?'
Don't understand	'If I asked you "what are you most adroit at?" What would you say?'
Don't understand	'If I asked you "where is the querulous cat?" What would you say?'
That's wrong	'If I said that you are a 2-year-old [BOY/GIRL], what would you say?'
That's wrong	'If I said that you live in Australia what would you say?

Each child was introduced to the ground rules in the same order (DU, DK, TW) and given two training questions which were pre-empted as 'practices'. If the child answered correctly on the first attempt (e.g. responds with "I don't know" to a DK question) then they were positively reinforced (e.g. "good job for telling me that you don't know the answer to the question") then progressed to the next question. If the child did not respond correctly on the first attempt (e.g. attempts to answer a DU question without acknowledging the part of the question that is incomprehensible) then the interviewer reminded the child of the ground rule (e.g. 'Let's check that, do you really know what 'querulous' means?..... 'It's important today that you tell me when you don't understand what a word means ok?'), before progressing to the next question.