2007

The Examination of Fixed and Multi-Tier Source Monitoring Training with Children

Sean Curtis Cameron
Wilfrid Laurier University

Your file Votre référence
ISBN: 978-0-494-32261-1
Our file Notre référence
ISBN: 978-0-494-32261-1

The Examination of Fixed and Multi-Tier
Source Monitoring Training with Children.

By

Sean Curtis Cameron
Bachelor of Science, Trent University, 1998
Master of Arts in Psychology, Wilfrid Laurier University, 2003
Bachelor of Education, University of Western Ontario, 2006

Doctoral Dissertation

Submitted to the Department of Psychology in partial fulfillment of the requirements
for Doctor of Philosophy in Psychology

Wilfrid Laurier University

Abstract

The current research consists of two studies examining children's source monitoring training. As previous research (e.g., Thierry & Spence, 2002; Poole & Lindsay, 2002) on source monitoring training is somewhat inconsistent, this research examined two different types of source training with 3-8 year old children. In Study 1, 131 children across two age ranges (3-4 and 7-8 years) were given comparable source training to that completed by Thierry and Spence (2002). General results indicated that the training benefited 7-8 year olds at two delay times, but only benefited younger children that met the established criterion in training. In Study 2, 136 children across 3 age ranges (3-4, 5-6 and 7-8 years) were given a newly developed multi-tier source monitoring training procedure. This training procedure was intended to scaffold ability and provide more individualized training to participants. General results showed no clear differences between training and no-training groups. However, when the results of Study 2 were compared to the results of Study 1, some evidence suggests that this training procedure may have inadvertently trained all older children in Study 2, thus washing out differences between training and non-training groups. Results are discussed in relation to the source monitoring framework, and forensic interviewing.

Acknowledgements

I would first like to thank my advisor Dr. Kim Roberts for her time and support throughout this long journey. Dr. Roberts has fostered my growth in the area of children's memory by having an amazing wealth of knowledge and experience in both the basic and applied fields of research. It is rare to see such an accomplished academic beautifully balance two fields of work, while remaining amiable and humble to students and peers. I would also like to thank my committee members Dr. Eileen Wood and Dr. Roger Buehler for their thoughtful reviews of this project.

I am indebted to the many children and families who have participated in this research, as well as the research assistants working within Dr. Robert's lab who have supported me as a team to complete this work. Without the support of these people, this work would not be possible.

Finally, I would like to thank my daughter Brianna Leigh Cameron, who was born during the completion of this project. On many days, her smiling face brought me the strength to keep working, and persevere.

Table of Contents

List of Tables

There is considerable interest in understanding the underlying factors that contribute to errors in children's eyewitness testimony. As children's eyewitness testimony is often the only evidence in cases where abuse has been alleged, the preservation of this is crucial. Both social processes (e.g., the authority of the interviewer) and cognitive processes (e.g., memory strength) have been studied. One cognitive process that has been repeatedly implicated as having an impact on the accuracy of children's eyewitness testimony is the ability to monitor the sources of one's memories (Poole & Lindsay, 1995; Powell, Roberts, Ceci & Hembrooke, 1999; Thierry, Spence, & Memon, 2001). Confusing the sources of memories leads children to inaccurately report information from a television program as if it happened in real-life (Roberts & Blades, 1998, 1999; Thierry et al., 2001; Thierry & Spence, 2002), or claim that false information suggested by an interviewer actually occurred during a witnessed event (e.g., Poole & Lindsay, 1995). The possibility that child witnesses can include such inaccuracies in their eyewitness reports is of great concern to forensic investigators given that cases are often built around children's testimony. The presence of such source confusions in children's reports has recently led researchers to examine whether children can be trained to monitor the sources of their memories and thus allow children to edit out inaccuracies in their eyewitness reports. The discovery and validation of a source-monitoring training technique would give forensic investigators an invaluable tool to extract reliable testimony from child witnesses.

Researchers are far from able to provide forensic investigators with a reliable and valid source-monitoring training technique. First, there are only a few published studies on source-monitoring training and more replication is needed before researchers can be confident about the reliability of such procedures, especially with regard to developmental differences in source-monitoring training. Second, the results that have been reported are inconsistent. For example, one procedure benefited 3- to 4-year-olds' source monitoring (Thierry & Spence, 2002), while another did not (Poole & Lindsay, 2002). It is important theoretically and practically to determine what might underlie differences in results. Understanding underlying factors will reveal how source monitoring develops and what prerequisite skills are needed for increasingly complex source judgments. There are marked developmental shifts in source-monitoring skills during the preschool years (see Roberts, 2002, for a review) and so there may also be developmental shifts in children's "readiness" to benefit from source-monitoring training, and developmental differences in the efficacy of different training procedures. Further, it is important to elucidate whether procedural differences affect children's success after training or whether young children lack the conceptual prerequisites (e.g., because of developmental limitations in prefrontal function, Drummey & Newcombe, 2002) to benefit from training at all.

Across two large studies, this research examines the idea of training children to report more accurate memories at the time of recall. In order to lay the groundwork for this project, I will begin with a general background of children's memories in relation to eyewitness testimony. Following this, two more specific types of memory theories will be discussed: source monitoring theory and fuzzy trace theory, and how these relate to the development of memory skills in children. Finally, the notion of training children in source monitoring (at the time of recall) will be discussed in relation to current research in this area (including the current studies). This research has specific implications for developing theories of children's memory, as well as indirect implications for forensic investigators who interview children.

Children's eyewitness testimony: A brief history

In recent decades, several paradigm changes have taken place in relation to how courtrooms view children as reliable witnesses. Traditionally, children were viewed as having 'fragile memories' and were not legally allowed to testify in courtrooms about their experiences; for fear of their memories being inaccurate (Ceci & Bruck, 1995; Ceci & Bruck, 1993). In the early 1980's, a shift developed in how forensic investigators extracted testimony from children. The premise was that children who have been abused, often need encouragement in order to disclose these atrocities (Ceci & Bruck, 1995). Thus, highly suggestive interviewing techniques were accepted as reliable by courtrooms for many years before researchers could uncover the danger in using these. As discussed in *Jeopardy in the Courtroom* (Ceci & Bruck, 1995), several high-profile cases of child sexual abuse in the United States reinstated the debate over the reliability of children's eyewitness statements. Following quite similar patterns, the "Little Rascals Daycare" and the "Wee Care Daycare" cases were two high-profile trials that substantially changed (in most states) how children are treated forensically when there are allegations of abuse (Ceci & Bruck, 1995). These cases involved highly sensationalized eyewitness accounts from children, which were extracted after highly suggestive and repeated interviewing from investigators (see Ceci & Bruck, 1995 for full review of cases). In Canada, courts were traditionally conservative in allowing the testimony of children, but recent changes have made children's eyewitness testimony a more vital component of criminal proceedings (see Bala, 1999 for a review).

Cases such as the "Little Rascals Daycare" scandal spawned a flood of research investigating the effects of suggestive interviewing, and a reconsideration of current 'best practices' used in the field. For instance, although at that time, it was common practice to use anatomically correct dolls for child forensic interviews, it is now generally viewed as a suggestive technique which is not recommended (Poole & Lamb, 1998).

After the scandals of the 1980's, much research in the area of children's eyewitness testimony initially examined 'contamination factors' in relation to interviewing children (i.e., what can we do wrong while interviewing children). However, more recently, researchers have begun to examine how we might go further and actually assist children in extracting more accurate memories at the time of recall. For example, recent research has examined numerous variables in relation to post-event eyewitness accuracy such as: the use of assistive aids in interviewing children (e.g., Brown, Pipe, Lewis, Lamb, & Orbach, 2007; Thierry, Lamb, Orbach, & Pipe, 2005), giving feedback about eyewitness identification to witnesses (e.g., Hafstad, Memon, & Robert, 2004), giving repeated interviews (La Rooy, Pipe, & Murray, 2005) and repeated events in relation to accuracy (Roberts & Powell, 2006) to name a few. In order to increase the accuracy of children's eyewitness testimony at the time of recall, researchers must examine the cognitive processes that are occurring at this time, rather than simply focusing on what was encoded at the time of the event.

Two main factors have been implicated for the increased amounts of errors in children's eyewitness accounts. These are generally referred to as 'social' and 'cognitive' factors. The first (which was focused on heavily after the sexual abuse scandals of the 80's), involves *social factors* that might have affected the child during interviewing (i.e., pressures, repeated questioning, etc.). It is now well established that children are more susceptible to social pressure in questioning than adults, even though this pressure might be subtle (Poole & Lamb, 1998). Specifically, children live in a world where they strive to please adults. At home and school, children are constantly asked to give the 'right' answers. Thus, they often assume that adults know these answers. Unfortunately though, in cases of alleged abuse, adult investigators do not know what has happened. When investigators repeat questions, or show any doubt in a child's answer, this can be construed by the child as an 'incorrect' answer; much like how a teacher or parent might repeat a question to a child if their first answer is incorrect (Poole & Lamb, 1998). Thus, children and adults can be susceptible to pressures of social demands during forensic interviews, and may develop answers that they believe the interviewer/s want to hear. For this reason, repeated interviewing, asking specific and leading questions (containing information the person has not already mentioned themselves), and repeating questions in forensic interviews are all considered practices which should only be used when absolutely necessary as they may increase the likelihood of a child making errors of commission (see Poole & Lamb, 1998).

The second factor which greatly affects the accuracy of eyewitness reports in children is the child's level of cognitive development. As memory decisions and recall are assisted with executive functioning (the controlling of major cognitive processes and thoughts) and frontal lobe development (see Johnson, Hashtroudi, & Lindsay, 1993 for a review), the development of cognitive functioning is a fundamental determinant in a child's ability to make accurate source judgements, and thus, recall accurate memories.

This issue is discussed in greater detail in relation to the development of children's source monitoring below.

How and what we ask of children during an interview is important in determining the level of detail and accuracy of their reports. For instance, recent research has shown that forced choice (yes/no) and recognition questions ("Was he wearing a red shirt?") can be difficult for younger children, as they may show acquiescence even when a suggested response is incorrect (Myers, Gramzow, Ornstein, Wagner, Gordon & Baker-Lynne, 2003). Further, children and adults tend to overestimate the accuracy of their responses, and are often confident of their answers, even when these are incorrect (Roebers, 2002).

The delay between event and interview also has a significant effect on memory recall. Much like any memory process, as time passes, memory degrades, and thus, so too does accuracy (Peterson, 2005). However, delay is not the only factor that has been implicated in affecting the accuracy of children's memories. Repeated questioning and repeated interviews have been found to have strong effects on how children later recall events and details in that they may incorporate post-event suggestions into their recall of the original memories (Pipe, Sutherland, Webster, Jones, & La Rooy, 2004; Poole & Lindsay, 2002; Powell & Roberts, 2002).

Most research that followed the high-profile sexual abuse scandals of the 80's was focused on identifying how (or if) we could extract accurate eyewitness accounts from very young children (e.g., less than 5 years). As language and cognitive development are happening at an exponential rate during this time, it seemed challenging to identify at 'what age' children could be reliable witnesses (if any). Much of the controversy of the

80's stemmed from children making errors of commission; that is, saying things that did not really happen. These errors are forensically more dangerous than errors of omission (not saying something that did happen), and thus, gained a lot of research attention. These errors are more dangerous as they have the potential (and have done so in recent history) to implicate potentially innocent people. One key theoretical notion that came of this research was the question of whether children could accurately identify 'memory source'. Researchers initially referred to this concept as 'reality monitoring'; one's attempt at distinguishing between things we imagined, versus things that have actually occurred (Johnson, Hashtroudi, & Lindsay, 1993). However, it was later suggested that the notion of recalling specific memory sources went well beyond the simplistic view of reality monitoring. There were many other 'sources' to choose from than merely "imagined" and "real".

The background of children's source monitoring

Source monitoring has been generally referred to as identifying *the origin of* a memory (i.e., the source). According to the source-monitoring framework (Johnson et al., 1993; Lindsay & Johnson, 2000) there are two main ways that the sources of memories can be accurately identified, and these two processes differ in the level of automaticity involved. The first process is automatic and involves an examination of the qualitative characteristics of memories as they are retrieved. Memories of experienced events contain more perceptual, contextual, affective and semantic information than do memories of non-experienced events. In addition, non-experienced events contain more information about the cognitive operations that took place at the time of the event. A second way that memory sources can be identified is more effortful and systematic.

Strategic processes can be deliberately invoked to determine the origin of memories, such as reasoning ("I couldn't have met you in Chicago because I've never been there; I must have imagined it"), noting relations, and retrieving supporting memories (Johnson et al., 1993). When we ask people to 'think hard' about where they have heard or seen something, it is likely that they are using strategic processes in order to make an attribution about the source of the memory. Although this attribution process may seem straightforward on the surface, there are many source decisions which are difficult to make, as we are deciding between two or more equally plausible source options (e.g., did I hear that on the radio or the evening news?). Source decisions tend to get more difficult: as the similarity of plausible sources increases (qualitative characteristics such as visual, audio and contextual cues), as the delay between event and memory recall increases, and familiarity of sources increases (Johnson et al., 1993).

As mentioned in Roberts (2000), two main factors affect the accuracy of our source judgements: the quality of the encoded memory, and secondly, the quality of the attribution process made at the time of judgement. As with any memory facet, time can have a seriously detrimental effect on the quality of the encoded memory. As time passes, so to does the quality of the distinct characteristics of our memories which make them distinguishable from one another. For instance, although you may remember very well what you had for dinner last night, in one month's time, this detail may be extremely difficult to separate from the many other meals that are plausible options in memory. Not surprisingly, research has reliably shown that time delay (between event and recall) has a negative correlation with source accuracy (Thierry, Spence, & Memon, 2000; also see Roberts, 2000 for a review).

Although we may not be able to control the encoding of the original memory source or the decay of this, we can however, control (to some extent) our attribution processes at the time of recall. In this sense, a memory source can be compared to a modern encoded DVD movie. On the disk, there are many features which distinguish that disk from other movies (such as the audio, the video, the titles and movie menus). However, there are also features which are similar to other DVD movies (such as actors, references to movie companies, characters etc.). When we are asked to distinguish one movie from another, we must pay attention to the features on the disk which are different across movies, and thus, will help us to make accurate source judgements. Thus, asking the question "What James Bond movie did I watch last?" might be an easy question if you have only ever seen one James Bond movie. In such an instance, this movie shares very few qualitative characteristics (actor, characters, theme etc.) with other movies that might be plausible options. However, this question becomes a difficult one if you have a collection of James Bond DVD's at home containing the same actor, same character (James Bond), same director and same movie company. In making such an attribution, you must look for source differences, not similarities. Without encoding these features which differ across sources, our source judgements may be reduced to guessing.

From an applied perspective, source monitoring decisions can play important roles in courtroom decisions and outcomes. For example, if a child is asked "Did he touch you on Wednesday or Thursday?", a child may be forced to make a critical source judgement which may determine the outcome of the case. If an accused person has a solid alibi on Thursday, and the child makes an incorrect source decision, and attributes the memory to this day, the entire eyewitness statement from the child might be called into question (even if the details are correct, but source decision was not).

When we make effortful source attributions, we are using cognitive operations which depend upon the qualitative differences we can differentiate across our memories. The more differences exist, the easier the decision becomes (Johnson et al., 1993). However, not all differences are created equal in source decisions. Differences that help us attribute source must be *source specific*, and thus, must be ones that give us additional information about our source decision. For instance, remembering that James Bond drove one car versus another in two different movies will not (by itself) help you to identify which movie you last watched. However, knowing that you normally watch James Bond movies as they are released will help you to further identify that although the character (James Bond) is the same across DVD's, the actors who play that character are not. Thus, a more recent movie is likely to contain a younger actor. In making this effortful process, you have narrowed the source decision down considerably by identifying a difference across sources that is specific to the time frame of the film's release (and thus, when you might have watched it). As mentioned by Roberts (2000), by manipulating the qualitative characteristics that differentiates sources, experimenters can make source attributions more difficult. In the previous example, this would be similar to asking the question "Which of the James Bond movies containing Roger Moore have you watched most recently?". In this question, we have removed a key qualitative characteristic that differentiates the movies based on the dimension of time, and in doing so, made the question more difficult – even though the source choices have been reduced substantially (i.e., there are actually fewer movies to choose from).

Fuzzy trace theory

Although the current research focuses on source monitoring theory, it is noteworthy to acknowledge that an alternative view of memory exists in relation to recalling specific source details. Fuzzy trace theory suggests that instead of a decision process (or attribution) about where a memory has come from (as source monitoring theory suggests), we 'pull' either a verbatim or gist memory to decide (Brainerd & Reyna, 1990). Accuracy on that question will be determined by which of these traces is accessed. Verbatim memories are memories of the detailed event, which retain 'source information'. Fuzzy trace theory suggests that verbatim memories decay quickly with time, and younger children are less likely to use them (Brainerd & Reyna, 1990; Brainerd & Gordon, 1994). Gist memories however, are more resilient, and last longer (Connolly & Price, 2006). These are memories for the 'general script' of the event (i.e., how things usually go in similar situations) – rather than the details (which may help identify source information).

Fuzzy trace theory suggests that both gist and verbatim memories are encoded in parallel (Brainerd & Reyna, 1990). When a question is asked, we access either one or the other depending on the quality and ease of accessibility – and this defines the accuracy of our memories. Thus, information which is "gist-consistent" (i.e., information that fits the general script of events), is particularly difficult to reject if we are merely using our gist memories to do so (Roberts & Powell, 2006; Thierry, Spence, & Memon, 2000). According to fuzzy trace theory, if we have plausible or "gist-consistent" information, these questions will be particularly difficult for younger children or after a long delay or after repeated events (Brainerd & Reyna, 1990). In summary, fuzzy trace theory suggests that over time or greater delays, people tend to access these gist memories more frequently (as the verbatim memory has decayed). Similarly, younger children tend to rely more on gist memories, and thus, have a greater difficulty with source-specific questions. However, fuzzy trace theory suggests that no 'effortful processing' or 'decision process' is made when a source memory is accessed. Just as fuzzy trace suggests that age is an important factor in determining memory accuracy, source monitoring theory suggests that as source attributions are directly related to executive function, age and cognitive development are key factors in determining source monitoring accuracy (Johnson et al., 1993; Roberts, 2000/2002).

The development of source monitoring

The finding that children (i.e., less than 10 years of age) are generally less accurate at source decisions than adults is well established in research (see Lindsay, Johnson, & Kwon, 1991). Although age is a key factor in determining source monitoring accuracy, many other factors (discussed above) also contribute to this.

The greater the overlap between the qualitative characteristics of events, the more likely children are to confuse them (e.g., Lindsay et al., 1991). Thus, children sometimes find it hard to discriminate between memories of "real-life" events and television because memories of both events contain perceptual information (e.g., visual and auditory information) and both events can produce affective reactions (Roberts & Blades, 1998). In studies demonstrating confusions between memories of television and real-life events, young children are particularly prone to such confusions and these confusions decrease with age (Roberts & Blades, 1998, 1999; Thierry et al., 2001; Thierry & Spence, 2002).

Johnson et al. (1993) suggest, automatic processing is a more basic and 'heuristic' strategy; often based on quick source judgments made by considering qualitative differences across memory sources. In contrast, Johnson et al. (1993) suggest that source decisions can be made more effortful by considering other factors that may determine source accuracy (e.g., "I know there was a bowl on the table, so crushing the crackers must have happened in real life"). The developmental evidence suggests that children become efficient at automatic source processing before strategic source processing. First, young children aged 3- to 4-years show more evidence of accurate source-monitoring decisions when they are tested behaviourally rather than verbally. For example, Roberts and Blades (1995) found that 4-year-olds could distinguish between memories of what they had done and what they had pretended to do in a hiding task when given a non-verbal test (i.e., they were asked to show only those places where they had *really* hidden counters) but not when given a verbal test ("did you really hide a counter under the tin?"). It is likely that this "verbal lag" in making source decisions may be due to a developmental difference in children's ability to make more effortful processing towards source decisions, but this remains unconfirmed in research. Second, in other research, young children can act on knowledge gleaned from conflicting sources of information but show no explicit awareness of where the information originated. For example, when young children looked at an object hidden in a box, they were more likely to claim the object was the one they saw than to change their response to what an experimenter who did not look in the box claimed it was (Whitcombe & Robinson, 2000).

Many other factors have been suggested as being linked to source monitoring development in children. For example, the development of 'Theory of Mind' has been suggested as an important cognitive milestone in developing an awareness of source monitoring (i.e., I can now take another's perspective, therefore I can consider another source; Welsch-Ross, Diecidue, & Miller, 1997). Similarly, resistance to misinformation and conflicting mental representations have also been linked to source monitoring development (Ackil & Zaragoza, 1995). The debate however, of what best predicts source monitoring accuracy in children is well beyond the scope of the current research. All of these factors do have a common link though, in that they are heavily dependent upon the development of executive function.

Training children in source monitoring

Although orienting children to the sources of their memories when their memories are being tested sometimes increases accuracy, the efficacy of source orientation varies by age. While source orientation reduced source errors in the reports of children aged 5- to 6- and older (Ackil & Zaragoza, 1995; Leichtman, Morse, Dixon, & Spiegel, 2000; Poole & Lindsay, 2001), such benefits were seen in 3- to 4-year-olds' reports in some studies (Giles et al., 2002; Thierry et al., 2001) but not others (Leichtman et al., 2000; Poole & Lindsay, 2001). It is possible that benefits of source-monitoring training (SMT) with young children can be consistently seen only when more intensive training is offered (i.e., children are trained to a minimum level of competency). This could be because of younger children's inability to consistently engage in strategic source monitoring. It may be that young children need to be trained until the source decision processes become automatized.

Recently, several studies have used more explicit source training to examine if children can benefit when recalling similar events. Explicit source training not only orients children to source (e.g., "think about where you learned things from"), but also corrects children on their accuracy (e.g., "actually it was in the *video* you saw that"). Importantly, these studies train children in source monitoring using non-target sources to see whether children can then generalize the training when asked to remember the target sources.

In a follow-up of their 2001 study, Poole and Lindsay (2002) used this explicit "feedback" training to examine if children aged 3- to 8- years old could benefit from source-monitoring training. Children were asked source questions about actions that were just performed or were merely described without being performed, and feedback was given (e.g., "That's right, I really did sharpen my pencil, you know that because you saw me do it"). Poole and Lindsay (2002) essentially replicated the findings of their 2001 study, by finding that explicit feedback benefited older children (aged 7-8 years), but had no benefit for younger children (aged 3-6 years). Poole and Lindsay (2002) suggested that perhaps younger children did not have the metacognitive ability to learn and generalize source-monitoring training to highly similar (and thus difficult) source-monitoring decisions (consistent with their 2001 findings). By metacognitive ability, Pool and Lindsay (2002) meant that children may not have sufficient development of their executive function; a system found the frontal lobe which is responsible for planning and deliberative decision making (something that SMT would require if an effortful attribution is to be made).

Thierry and Spence (2002) also conducted a follow-up to their earlier study (Thierry et al., 2001) using more explicit source-monitoring training procedures (with feedback using non-target items). Children in this study were given accuracy feedback on questions asked about a puppet show or TV show they had just watched (e.g., "Did you see Billy crown himself King?; In real life or on TV?"). Thierry and Spence (2002) trained 3- to 4-year-olds to a criterion level of training (until the child got four consecutive recognition and source questions correct) and then questioned children about two "Mrs. Science" lessons (one presented live, and the other on TV). They found that the source-monitoring training procedure significantly increased the young children's accuracy (in relation to a recognition-trained control group) when asked both yes/no questions about items that were and were not present in the target sources and when asked non-misleading open-ended questions.

There are many methodological differences between Poole and Lindsay's (2002) and Thierry and Spence's (2002) studies that may explain the discrepant results, such as the target events, the delays between the target events and the memory tests, and the frequency of source presentation. As both groups contained controls who did not receive any source training, however, these factors might affect the magnitude of training effects but not necessarily the presence or absence of training effects. One of the more striking and theoretically interesting differences between the studies refers to the intensity of the training procedure. Poole and Lindsay provided children with a set of three training questions with feedback, whereas Thierry and Spence continued to provide questions and feedback until the children in their study had reached a criterion level of training (presumably indexing that the children had acquired the particular source judgment required for this task). Specifically, Thierry and Spence (2002) required children to be correct on four consecutive source questions. Poole and Lindsay (2002) did not focus on criterion training but reported that 21% of the 3- to 4-year-olds and 7% of the 7- to 8-

year-olds were not consistently correct on the three source-monitoring training questions that probed details that were merely described. Thus, it is possible that the 3- to 4-year-olds in Poole and Lindsay's study were not trained to the same level as the children in Thierry and Spence's study. Ensuring children reach a training criterion (as an indication that they are actually "trained") may be especially important with young children who show significant individual differences in metacognitive and source-monitoring abilities (see Leichtman et al., 2000; Quas, Qin, Schaaf, & Goodman, 1997).

The differences in results could also reflect developmental differences in readiness to benefit from source training given that young children are capable of implicit source monitoring but struggle with more strategic and reflective source processing (e.g., Whitcombe & Robinson, 2000). Perhaps the older children in Poole and Lindsay's study were already capable of monitoring the sources but needed to be instructed to do so. The younger children, on the other hand, were unlikely to possess such sophisticated source-monitoring skills. Younger children may be able to complete source-monitoring tasks successfully only if they are trained until they have perfected the source-monitoring requirement (i.e., instructions to engage in source monitoring may not be sufficient). Alternatively, it could be that Thierry and Spence's results were the result of a Type I error and younger children simply do not have the cognitive prerequisites to improve their source monitoring. As these findings have yet to be replicated in another lab, further research is necessary to validate the successful training of younger children.

Although results of source monitoring training studies have been somewhat inconsistent, research in this area has developed considerably, and thus, methodologies have changed substantially. For instance, although all of the studies found *some* benefit of source training, only Leichtman et.al., (2000), Thierry et al., (2001), Thierry and Spence, (2002), and Poole and Lindsay, (2002) used a non-training control group to compare training and non-training groups. Further, it was only Poole and Lindsay (2002) and Thierry and Spence (2002) who used non-target items for testing the effectiveness of their SMT procedures. This detail may seem trivial, but is a key element of Brown and Kane's (1988) transfer of training (discussed in more detail below). Training studies should 'train' participants using one set of variables and items, and then 'test' participants using a different set of variables (to dissect the issue of transfer versus immediate training). Thus, the current studies have mainly been based on Thierry and Spence (2002) and Poole and Lindsay (2002), as these are the most recent and most methodologically rigorous source monitoring training studies to date.

To address these differences in findings, we directly compared several different types of source monitoring training procedures comparable to those used in previous studies (set amount of trials versus training to criterion) while controlling other methodological differences (e.g., events, delays). Children aged 3- to 4- and 7- to 8- (i.e., the same age groups used in previous research showing developmental differences) participated in a lesson about digestion and watched a video recording of a lesson, and were tested 2-3 days or 7-10 days after viewing the target sources. A different set of events to that used by Poole and Lindsay (2002) and Thierry and Spence (2002) was chosen to rule out the possibility that differences were due to differences with the specific target events in these studies. Only older children (7-8 years) were run at the longer delay because younger children's scores at the 2-3 day delay suggested that they would find the task too difficult at a longer delay. It was hypothesized that older children (7-8 years) would show training effects regardless of whether they received a set amount of training or were trained to criterion. This hypothesis is consistent with Poole and Lindsay's results (2001, 2002), who found that only general source orientation was sufficient to find source-monitoring training effects with older children. We hypothesized that obtaining training effects with younger children (3-4 years) would be more difficult, however, due to the metacognitive abilities required to generalize source-monitoring training to a target task. As our target sources were quite similar (and thus the source task difficult), it was hypothesized that if any training effects were found with the youngest children, this would only be within the criterion-training condition (consistent with Thierry et al., 2001; Thierry & Spence, 2002).

Although the current research is not a strategy development study, it is helpful to consider this literature in relation to source monitoring training. The aforementioned training paradigms can be viewed in relation to Flavell's (1970) model of using strategic processes to benefit memory. Flavell (1970) suggested that *mediational deficiencies* are present when a child simply does not have the cognitive prerequisites to benefit from strategy training. A more developed child may however, have the ability to produce and use a strategy, but may show little or no benefit from its use beyond a training session. This child would be considered *utilization deficient* as they are not utilizing the strategy properly. If however, a child can be trained to use a strategy efficiently and benefit from its use (but did not use the strategy before training); they would be considered *production deficient.* Table 1 compares the differences between Flavell's (1970) stages of strategy use, in relation to an extensive review of these in published memory training studies completed by Bjorklund, Miller, Coyle, and Slawinski (1997).

Table 1. Stages of Memory Strategy Usage as identified by Bjorklund et al. (1997)

Stage of Strategy Usage	Show high levels of task performance in training (i.e., can use a strategy)	Show benefit from strategy use in post-training task (i.e., increased performance)
Ineffective training	No	No
Mediational deficiency	No	No
Utilization deficiency	Yes	No
Production deficiency	Yes	Yes

Recent research on source monitoring training (SMT) has suggested that children who benefit from SMT (over a non-training control group), may be working from a production deficiency, and that SMT (when done intensely enough) might benefit children's source accuracy (Thierry & Spence, 2002). There is disagreement however, as to whether children who *do not* benefit from SMT are showing a mediational deficiency, or merely a utilization deficiency. If children are in the mediational deficiency phase, no amount of training would benefit their ability to monitor sources, as they simply do not have the cognitive prerequisites yet. Thierry and Spence (2002) have suggested that because 3-4 year old children in their 2002 study reached a criterion level of source monitoring during training, that when younger children do not benefit from training (as was the case in Thierry et al., 2001), they may simply be experiencing utilization deficiency rather than mediational deficiency. Thierry and Spence (2002) suggest that a progression from utilization deficiency to production deficiency may be made with intensive source training. As the Thierry and Spence (2002) study included a criterion level of training (a minimum standard children needed to pass to be considered "trained"), this study was able to show that these children could use a strategy to benefit their source monitoring, meaning they were not in the mediation deficiency stage. Poole and Lindsay (2001/2002) conducted two studies examining 3-8 year old children (see Table 2). In both studies, they found that SMT did not significantly benefit younger children (3-4 years) as it did in the Thierry and Spence (2002) study. Poole and Lindsay (2002) concluded that young children (3-4 years) simply did not have the cognitive prerequisites to benefit from SMT, and thus were in a stage of mediation deficiency. However, Bjorklund et al. (1997) suggested that to differentiate between mediational deficiency and ineffective training, a study must test children to some level of criterion prior to their recall task (as Thierry & Spence, 2002 did). Bjorklund et al. (1997) also suggested that increased amounts of training may decrease utilization deficiencies, and that utilization deficiencies may be dependent on the difficulty of the test and appropriateness of the test to the age of the child.

Table 2. Summary of published SMT studies with children.

Study	Ages Used	Did older children benefit from SMT?	Did younger children benefit from SMT?	Was SM ability compared to a non-training control?	Was SMT explicit with feedback?	Was a criterion level of training used?
Leichtman et. al., 2000	3-6 years	Yes (5-6 years)	No (3-4 years)	Yes	No	No
Poole & Lindsay, 2001	3-8 years	Yes (5-8 years)	No (3-4 years)	No	No	No
Thierry et al., 2001	3-6 years	Yes (5-6 years)	No (3-4 years)	Yes	No	No
Giles et. al., 2002	3-5 years	N/A	Yes (3-5 years)	No	No	No
Poole & Lindsay, 2002	3-8 years	Yes (7-8 years)	No (3-6 years)	Yes	Yes	No
Thierry & Spence, 2002	3-4 years	N/A	Yes (3-4 years)	Yes	Yes	Yes

Research in the SMT area must then, adequately train children to ensure that they are at least within the utilization deficiency stage. Insufficient training in SMT studies confounds the understanding of whether children who did not benefit from training simply were not trained enough (and are utilization deficient), or did not have the cognitive prerequisites to benefit from it (and are mediational deficient). Only one study to date (Thierry & Spence, 2002) has differentiated between utilization deficiency and mediational deficiency by establishing some level of criterion in training. In sum, even young children aged 3- to 4-years show some implicit source-monitoring abilities but do not appear to be able to engage either spontaneously or strategically in more reflective source processing. Thierry and Spence (2002) suggest that this lack of spontaneous, strategic source processing in younger children may be the result of a production deficiency (see Table 1; Flavell, 1970). If this were the case, these children's memories should be facilitated with training. There are dramatic developments in the ability to explicitly reason about sources after the age of 4, however, though children still make more source-monitoring errors than adults until about age 10 (see Roberts, 2002, for a review).

It remains unclear at what point children have the cognitive prerequisites to benefit from source training, and thus make the move from being mediation deficient (cannot use the mediator to facilitate performance), to being production or utilization deficient (can use strategies when trained). It is also unclear at what point children who use strategies when trained (utilization deficiency), begin to benefit from their use on a memory task (production deficiency). To explore this notion, a study must establish some criterion level of performance in training (to show usage of strategy), and then test for generalized use of this strategy on a subsequent task. If a child cannot meet the training criterion, they are believed to be experiencing a mediational deficiency. If they can meet the training criterion, but do not benefit from training, they are believed to be utilization deficient. Finally, if they meet the established criterion *and benefit* from training on a subsequent task (more than a non-training control group), they are believed to be production deficient (see Table 1 above). Thus, it is important to have both an established criterion in training, as well as a subsequent generalization test in order to identify what stage of strategy usage a child may be at. A training study should also compare the post-training scores to a non-training control group in order to establish if benefits from training were merely children effectively using a previously learned strategy (i.e., they had the skill before training), or if they were actually in the production deficiency stage. As shown in Table 1, Thierry and Spence (2002) is the only published SMT study which used an established criterion, and compared results to a non-training control group. The established criterion in this study was to continue training until children got four consecutive questions correct. This study showed that children were not simply using a pre-existing skill (i.e., training benefited them), and that these children were at least in the utilization deficiency stage (as they passed the established criterion). Had this study shown no effects of training with younger children (which it did not), Thierry and Spence (2002) would have at least established that these children were not mediation deficient. Unfortunately, as no other SMT studies have included any criterion of performance in training, these studies are not clear as to whether children are showing mediational deficiencies or utilization deficiencies (i.e., can they even do the task?).

Brown and Kane (1988) might suggest that to examine the effectiveness of SMT and better understand any benefits a procedure might yield, a study should include both a non-training control group, as well as some minimum level of criterion performance in training. However, what is not yet clear, is whether a 'set amount' of training is sufficient to benefit children (as was the case in Poole and Lindsay, 2002), or whether they require a 'criterion amount' of training (as was the case in Thierry and Spence, 2002). Further, as all participants in the Thierry and Spence (2002) study met criterion, it is unclear if the notion of 'criterion' is important, or if these participants merely benefited from being asked several source questions.

In Study 1, we predicted that source training of any level would benefit older children (7-8 years), but that children in the youngest age group would need a criterion level of training if they are able to benefit from a source training procedure. Study 2 predicted that an individualized source training procedure would better help younger and older children to benefit from source monitoring training.

Study 1

Study 1 directly compared the issue of using a 'set amount' of training versus 'criterion' training, and compared these two training groups to a non-training control group. It was hypothesized that older children in the criterion or set-training condition would benefit from source training, but that any benefit of training to younger children might be limited to the criterion condition (comparable to the findings of Thierry & Spence 2002).

Method

Participants

One hundred thirty-one children were recruited from three local daycares and two elementary schools in a mid-sized North American city. Thirty-four 7- to 8-year-olds ranged in age from 84 to 105 months (M = 8 years, 0 months) and were interviewed 7- to 10-days after viewing the target events (M = 7.64 days, SD = .74). One 7- to 8- year old child was removed from the study due to being absent for the interview. Forty-six other 7- to 8-year-olds (M = 7 years, 6 months; ranging from 83-105 months) and 51 3- to 4-year olds (M = 4 years, 2 months, ranging from 36-59 months) were interviewed 2- to 3-days after viewing the target events (M = 2.31 days, SD = .46). Three of the younger children were absent for their interviews and thus were excluded.

At the beginning of the experiment (i.e., prior to participating in the training session), children in each Age x Delay group were randomly assigned to one of three conditions: either a recognition-training (control) group, or one of two source-training groups (set training, criterion training).

Parents gave informed consent and all children gave their assent to each session. Once the study was complete, children were given a toy and a video copy of the digestion lesson in which they participated. Schools and daycares were also given small financial donations for their participation.

Materials

Pre-developed scripts and question sheets were used for all sessions (see Appendix A for full list). Video cameras and audio recorders were used to record all sessions. Materials for the lessons included various materials that were purchased at educational stores and dollar stores fitting the 'theme' of the script (e.g., nature lesson, frog lesson, digestion lesson).

Procedure

Session 1- The Target Event

Children were informed that they would be learning about digestion and were subsequently escorted in groups of 2-4 to an empty classroom within the school by a male researcher. The topic of digestion was chosen because the topic is relatively unfamiliar to children (thus children are unlikely to correctly answer questions on the basis of general knowledge rather than episodic memory). The digestion lesson comprised a 5-minute "All about digestion" video, and a highly similar 5-minute "All about digestion" activity (referred to as the "real life game" to participants). The order of presentation was counterbalanced across conditions.

The activity involved interactive props, and experiments relating to learning about digestion (e.g., mixing crackers in a bowl and adding water). The video was previously filmed in the lab and featured a different male researcher. The video and activity only differed in 32 target details that were counterbalanced between the video and live activity and comprised actions and perceptual details typical of other memory research (e.g., confederate demonstrated gas fizzing in a pickle jar in the video but not the activity). All target items were clearly presented and verbally described to ensure that children paid attention to the target items. To ensure no confounding of specific target items with source modality, two different forms of the script were created by randomly assigning items to one of two versions (and subsequently two different videos). Participants were randomly shown either version 1 or version 2 of the digestion scripts. Children were told that they would be watching a video about digestion, and then doing a real life game about digestion (order counterbalanced; e.g., "First we are going to watch a video about digestion [point to TV], and then we will do a real-life game about digestion" [point to table]). All children were familiar with videos. The "video" and "real life game" labels were repeated throughout both the event and training (see below), so that children were clear about the labels for each source. After the completion of the digestion activity and video, the male researcher thanked the children and escorted them back to their classes. Children of all ages were engaged and attentive during both the video and live activity presentations and found the lessons enjoyable.

Session 2 – The Interview

After a 2-3 or 7-10 day delay, children were approached individually by a female research assistant. The children were asked if they wanted to learn about frogs, and once assent was given (all children gave assent), they were escorted to an empty room in the school. The children underwent the training portion of the interview (i.e., questions about the frog lesson), and subsequently the follow-up target interview about the digestion lesson.

Training phase (Frog lesson).

The children learned about frogs through the same source modalities as were used in the previous digestion lesson. Children watched a 3-minute "All about Frogs" video and a 3-minute activity about frogs (presentation order counterbalanced). Target items were randomly assigned to one of two counterbalanced forms, and each child was randomly given either Form 1 or Form 2 of the frog lessons. As with the digestion lesson, the frog activity and video were interactive and involved props and actions on the part of the child (e.g., placing frogs into an aquarium).

Once children had watched the frog video and live activity, the props used in the frog lesson were covered with a blanket and the training began. Source-monitoring training involved asking children a series of recognition-source question pairs and providing feedback on the accuracy of their responses (feedback was given for both accurate and inaccurate responses during training). Specifically, children were first asked a recognition question about what they had just witnessed (e.g., "Did a raccoon chase the frog?") and, following a correct 'yes' response (all children were correct), were asked a source question that probed whether the item occurred in the "video" or "real life game" (order of options counterbalanced). Training questions included a balanced of video and game specific training questions (i.e., some items occurred in the video, while others were in the source). All training questions were non-misleading in nature (i.e., all items asked about actually occurred). Feedback was given to correct responses to source questions by saying (for example), "That's right. A raccoon chased the frog in the *video*" and to incorrect responses by saying (for example) "That's a good guess but actually the raccoon chased the frog in the *real life game*". Motivational encouragement with further reference to source was given after the completion of the training by saying "It looks like you're good at thinking about where you learned things from".

Children in the *criterion-training* condition were trained (i.e., given recognition-source question pairs followed by feedback) until they reached a "criterion" level of training. Following Thierry and Spence's (2002) procedure, criterion was reached when children correctly answered four consecutive recognition-source question pairs, and training was terminated if criterion was not met after 20 recognition-source question pairs. Note, however, that children were not asked about "trick" questions during training (misleading questions) in contrast to the children in Thierry and Spence's study. Of the 20 children in the 7- to 8-year-old group, 16 met criterion after four trials, 2 after six trials, and 2 required 18 trials. Of the 19 3- to 4-year-olds, 4 met criterion after four trials, 10 required 6-15 trials, and 5 were given all 20 questions.

Children in the *set-training* condition were given the same source training, feedback, and motivational instructions as those in the criterion-training condition except that only a set amount of source training was given (rather than training until the children met criterion). Specifically, children were asked four recognition-source question pairs and then training was terminated. The same motivational encouragement about source memory was given as in the criterion-training condition.

Children in the *recognition,* or control, condition were asked four recognition questions without feedback and were then encouraged (i.e., "It looks like you're good at remembering things"). Note that no questions about source (video or activity) were asked, children did not receive any training in source identification, and the motivational instructions referred to general recognition rather than source identification.

The types of training questions that children were asked involved non-misleading forced choice recognition questions (e.g., "Did the frog catch a fly with his tongue?"), followed by a choice of source (order counterbalanced) if the child said "yes" to the source question (e.g., "Was that in the game or the video"). Details probed in the training questions were balanced across the two sources, and instantiations (i.e., a plausible option in the alternate source) were also balanced across questions.

Target phase (Digestion lesson).

Upon completing the training questions, the interviewer gave a naïve introduction to the digestion events (i.e., "I heard you met [researcher] last week. I wasn't there and don't know what happened. I'd like you to tell me only about the things you learned on that day with [researcher]"). The children were asked about each of the 32 target items in the digestion video and activity, thus, children's memories were probed for actions, perceptual details, and verbal statements. Four interview forms (two forms for each of the two digestion versions) were created so that the order of target questions was counterbalanced.

There were 16 yes/no *recognition* questions - half were non-misleading as in "Was there blue water?" (correct answer was 'yes'; 4 questions probed activity items, 4 probed video items) and the rest were misleading as in "Was there orange water?" (correct answer was 'no'; 4 questions probed activity items, 4 probed video items). Correct and incorrect 'yes' responses to the recognition questions were followed with *forced-choice source* questions as in "Was there blue water in the real-life game or the video?" (order of sources counterbalanced). The number of correct source identifications were divided by the number of forced-choice source questions asked. Question types and formats were similar to those used in previous source monitoring studies (e.g., Thierry & Spence, 2002; Poole & Lindsay, 2002). All items asked about in the study were developed ahead of time, and subsequently randomly assigned to a 'question type' (e.g., misleading, non-misleading). After the items and question types had been developed, two different forms of scripts were developed based on these items and question types.

Children were also asked 16 *combination* questions in which the detail and a source were combined. In eight of these questions (*misleading-source* questions), a detail was paired with the alternate source as in "In the real life game, did he drop a kidney on the floor?" (when a kidney was dropped on the floor but it was in the video and not the real-life game, as suggested in the question). These questions were predicted to be particularly difficult for children in that even if they recognize the detail as one that was in the digestion lesson, they may still be misled about its source (activity or video). The remaining eight questions (*misleading-detail* questions) asked about a detail that was not present as in "In the video, did someone pour coffee through the filter" (when no one poured coffee in either source). The correct answer to the combination questions was 'no' even though the misleading-source questions probed details that were present.

Children were given no feedback about the accuracy of their answers, as the interviewer was blind to which counterbalanced versions of the events children had actually seen. The target events and interview sessions were video- and audio-taped for later coding. As all responses were forced choice or 'yes/no', coding rules were straightforward and inter-rater reliability was 100% (based on independent coding of 30% of the interviews). "I don't know" responses were rare (accounting for less than 2% of responses) and so were not analyzed.

Results

In the first section, we analyzed training effects in the 7- to 8-year-olds only. This analysis allowed a comparison of training effects across two different delays, that is, when training was conducted shortly after (2-3 days) or a while after (7-10 days) viewing the target sources. In the second section, we analyzed age differences in training effects at the 2-3 day delay. Tests were two-tailed unless otherwise mentioned, and LSD post-hoc tests were used to analyze the main effects of training.

Training effects with 7-8 year-olds across two delays

Responses to the recognition questions were analyzed first. The proportion of accurate responses were entered into a 3 (Training condition: criterion, set, control) x 2 (Delay: 2-3 days, 7-10 days) x 2 (Question type: non-misleading, misleading) analysis of variance (ANOVA) with repeated measures on the last factor. There were effects of training, $F(2, 70) = 4.40, p = .02$, delay, $F(1, 70) = 8.91, p = .01$, a marginal effect of question type, $F(1, 70) = 2.83$, $p = .01$, and a marginal interaction between question type and delay, $F(1, 70) = 3.63$, $p = .06$. Regarding training, children in the criterion group gave a greater number of accurate responses than did children in the set and control conditions (see Table 3 below for means and standard deviations). One exception to this finding is that although the Question type x Condition x Delay interaction was not significant $F(2,70) = 1.05$, $p = n.s.$, it is clear (see Table 3 below) that children in the criterion condition at the shorter delay were not more accurate on non-misleading recognition questions than other groups. Regarding delay, responses to misleading questions were more accurate at the shorter delay than the longer delay (see Table 3 below for means), $F(1, 74) = 12.98, p = .01$. Regarding question type, a greater number of accurate responses were given to the non-misleading than misleading questions, but only at the longer delay (*M*s = .64, .52, *SD*s = .19, .21, respectively), $t(30) = 2.28, p = .03$.

Table 3: Means and standard deviations (in brackets) of misleading and non-misleading recognition scores for older children across delays (7-10, 2-3 days) and conditions.

		Non-misleading Recognition	Misleading Recognition
2-3 days			
	Control	.71 (.15)	.61 (.18)
	Set	.61 (.14)	.63 (.21)
	Criterion	.67 (.19)	.78 (.19)
7-10 days			
	Control	.61 (.17)	.50 (.19)
	Set	.66 (.22)	.44 (.25)
	Criterion	.67 (.19)	.64 (.15)

To analyze the source-monitoring scores, a 3 (Training condition: criterion, set, control) x 2 (Delay: 2-3 days, 7-10 days) multivariate analysis of variance (MANOVA) was run on the proportion correct responses to the three question types that included reference to source (i.e., the forced-choice source, misleading-detail, and misleading-source questions). There were multivariate effects of training, Wilks' $\lambda = .69$, $F(6, 136) = 4.54$, $p = .000$, and delay, Wilks' $\lambda = .84$, $F(3, 68) = 4.42$, $p = .01$. Box's test for homogeneity of variance was explored and was not violated with any multivariate test completed.

Specifically, significant training effects were obtained for the forced-choice source questions, $F(2, 70) = 3.60$, $p = .03$, and the misleading-source questions, $F(2, 70) = 7.83$, $p = .01$. As expected, children trained to criterion were more accurate when responding to the misleading-source questions than were children in the set-training and control conditions (*Ms*= .46, .30, .28, *SDs* = .18, .15, .16, respectively). Somewhat unexpectedly, children in the set-training condition were more accurate in response to the forced-choice source questions than children in the criterion and control conditions (*Ms* =

.74, .63, .60, *SDs* = .19, .13, .21, respectively).It is worth noting that all children in the criterion condition scored 50% or higher on non-misleading forced choice questions, whereas the mean for children in the set and control conditions began at 25% (control) and 38% (set). Thus, although the criterion-training condition mean did not significantly differ from that of the other groups, all children in the criterion group appeared to show training effects but not all children in the set-training condition did so.

There were significant univariate effects of delay in responses to the misleading-source questions, $F(1, 70) = 7.46, p = .01$, and the misleading-detail questions, $F(1, 70) = 10.68, p = .01$. As expected, children were less accurate at the longer delay of 7-10 days (*M*s = .28, .48, *SE*s = .03, .039, for misleading-source and misleading-detail questions, respectively) than the shorter delay of 2-3 days (*M*s = .38, .65, *SE*s = .024, .031, respectively). In sum, these analyses showed that children in the criterion condition were significantly higher (than the other two conditions) on the most difficult question types (combination misleading source), whereas the set training group scored significantly higher (than the other two conditions) on the forced choice source questions.

All children were included in the above MANOVA whether or not they met criterion in the criterion-training condition. In fact, only one child in the criterion-training condition did not meet criterion. The above MANOVA was repeated excluding this child and the results were identical across training conditions and delays: training effect, Wilks' $\lambda = .72, F(6, 135) = 4.24, p = .01$,; delay, Wilks' $\lambda = .81, F(3, 67) = 5.74, p = .01$. Finally, the MANOVA was repeated excluding the 6 children in the control condition and the 7 children in the set condition who did not meet criterion were excluded from the analysis ('criterion' for children in the control condition was considered to be four consecutive correct responses to the recognition questions only, Thierry & Spence, 2002). Once again, the main effects were identical: training, Wilks' λ = .70, $F(6, 123) = 3.88$, $p = .01$, delay, Wilks' λ = .78, $F(3, 55) = 5.43$, $p = .01$. In sum, source-monitoring training in the criterion condition benefited the 7- to 8-year-olds when they were asked the most difficult question types (combination misleading source). Similarly, children in the set training condition benefited when they were asked forced-choice non-misleading source questions. When the children who did not meet criterion were excluded from this sample, training effects remained unchanged. Similarly, when recognition scores were entered as a covariate, none of the training effects across any analyses changed.

Training effects with two age groups at a shorter delay

To analyze responses to the recognition questions, the proportion of accurate responses was entered into a 3 (Training group: criterion, set, control) x 2 (Age: 3-4 years, 7-8 years) x 2 (Question type: nonmisleading, misleading) ANOVA with repeated measures on the last factor. There was an effect of age, $F(1, 89) = 45.41$, $p = .01$, $\eta_p^2 = .34$, because the older children ($M = .67$, $SD = .26$) were more accurate than the younger children ($M = .52$, $SD = .30$). The type of training did not affect recognition.

To compare age differences in training effects on questions containing reference to source, a 3 (Training group: criterion, set, control) x 2 (Age: 3-4 years, 7-8 years) MANOVA was conducted on the proportion correct responses to the forced-choice source, misleading-source, and misleading-detail questions. There was an effect of age, Wilks' λ = .74, $F(3, 85) = 9.96$, $p = .000$, $\eta_p^2 = .26$, and an interaction between training condition and age, Wilks' λ = .85, $F(6, 170) = 2.32$, $p = .04$, $\eta_p^2 = .08$.

The univariate ANOVAs showed that there were age differences in responses to the forced-choice source questions, $F(1, 93) = 3.81, p = .05, \eta_p^2 = .04$, the misleading-source questions, $F(1, 93) = 12.08, p = .01, \eta_p^2 = .12$, and the misleading-detail questions, $F(1, 93) = 29.78, p = .01, \eta_p^2 = .26$. The 3- to 4-year-olds were less accurate than the 7- to 8-year-olds in all cases (3-4 year-olds: *M*s = .61, .25, .36, *SD*s = .26, .21, .29; 7-8 year-olds, *M*s = .70, .38, .65, *SD*s = .19, .17, .21, respectively).

The Training x Age interaction referred to responses to the forced-choice source questions, $F(2, 93) = 2.61, p = .04$ (1-tailed), $\eta_p^2 = .06$. The most accurate responses from the younger children were given by those in the criterion condition. Further, the younger children who were trained to criterion were as accurate as their older counterparts. This observation was confirmed through statistical analysis: Independent-samples *t*-tests with alpha lowered to .01 were conducted to compare age differences in the two training conditions. While there was an age difference in the set-training condition, $t(27) = -3.08$, $p = .01$, there were no age differences in the criterion-training condition, $t(30) = 0.44$, *ns*. In sum, younger children (3-4 years) were significantly less accurate than older children on all source questions. However, one exception was that children in the criterion condition were on average, as accurate as older children in the criterion condition. When t tests were conducted directly comparing the control group to the criterion group (in the 3-4 year old children), no significant differences were found for non-misleading forced choice source, combination misleading source or combination misleading detail questions: $t(24) = .44$, p = *n.s.*, $t(25) = .89$, p = *n.s.* and $t(25) = .81$, p = *n.s.* respectively.

Table 4: Means (and standard deviations) of target source question accuracy across age groups and conditions ($N = 94$).

Age x Condition x Question Type	Non-misleading forced choice source	Question type Combination misleading-source	Combination misleading-detail
3- to 4-year-olds			
Control	.61 (.33)	.26 (.25)	.39 (.29)
Set-training	.54 (.26)	.24 (.25)	.38 (.29)
Criterion-training	.64 (.22)	.24 (.18)	.36 (.31)
7- to 8-year-olds			
Control	.65 (.21)	.31 (.17)	.65 (.24)
Set-training	.79 (.16)	.33 (.16)	.68 (.20)
Criterion-training	.64 (.14)	.50 (.13)	.62 (.17)

The above MANOVA included the responses of all children in the criterion-training condition regardless of whether they met criterion or not. Thierry and Spence (2002) reported that all children met criterion in their study, yet this was not the case in our study. Thus, to better compare the results by Thierry and Spence and those in the current study, we repeated the MANOVA but excluded the responses of children in the criterion condition who did not meet criterion (four were 3- to 4-year-olds, and one was a 7- to 8-year-old). The results were identical to the MANOVA with the full sample (Age: Wilks' $\lambda = .75$, $F(3, 54) = 6.10$, $p = .01$, $\eta_p^2 = .25$; Training x Age: Wilks' $\lambda = .87$, $F(3, 54) = 2.75$, $p = .05$, $\eta_p^2 = .13$) except that a main effect of training also emerged, Wilks' $\lambda = .85$, $F(3, 54) = 3.13$, $p = .03$, $\eta_p^2 = .15$. Children in the criterion condition ($M = .39$, $SD = .20$) were more accurate than controls ($M = .29$, $SD = .19$) when answering the misleading-source questions, $F(1, 60) = 5.09$, $p = .03$, $\eta_p^2 = .08$. Planned t tests were conducted between the control group and the criterion group for 3-4 year old children – excluding those children in the criterion condition who did not meet criterion. Results showed that the criterion condition (M = .68, SD = .22) were more accurate than the control condition (M = .48, SD = .25) with marginal significance on non-misleading forced choice source questions $t(13) = 1.63$, $p = .13$. This low level of significance is likely an issue of power, as generally a larger cell size is required to find significance with younger children. No significance differences were found in younger children between the control and criterion groups on misleading combination questions $t(13) =$ .64, p = n.s., or on misleading detail combination questions $t(13) = 0$, p = *n.s.*. On all question types however, means for the criterion group were equal to, or above those of the control condition (*Ms* = .68, 26, .37) for criterion and (*Ms* = .48, .19, .37) for control respectively.

Finally, the MANOVA was run excluding the children who did not meet criterion in any condition ('criterion' for children in the control condition was considered to be four consecutive correct responses to the recognition questions only, Thierry & Spence, 2002). Within the 3- to 4-year-old group, only one child met criterion in the set-training condition, and so the set-training condition was not included in the MANOVA. In the control condition, 6/13 of the 3- to 4-year-olds and 4/19 of the 7- to 8-year-olds did not meet criterion and so (for the 3-to 4-year-olds and 7- to 8-year-olds, respectively), the MANOVA was conducted on 54% and 79% of responses from children in the control condition and 93% and 100% of responses from children in the criterion-training condition. Again, the results were identical to the full-sample MANOVA (Age, Wilks' λ = .77, $F(3, 35) = 3.38$, $p = .03$, $\eta_p^2 = .23$; Training x Age, Wilks' $\lambda = .80$, $F(3, 35) = 2.89$,

$p = .04$, $\eta_p^2 = .20$), except that a main effect of training also emerged, Wilks' $\lambda = .77$, $F(3, 35) = 3.46$, $p = .03$, $\eta_p^2 = .23$. Children in the criterion condition ($M = .37$, $SE = .04$) were marginally more accurate than controls ($M = .26$, $SE = .05$) when answering the combination misleading-source questions, $F(1, 37) = 3.44$, $p = .07$, $\eta_p^2 = .09$ overall, and younger children in the criterion condition were significantly more accurate than controls on non-misleading forced choice source questions, $F(1, 37) = 4.53$, $p = .04$, $\eta_p^2 = .11$ (see Table 5 below). When planned t tests were conducted comparing performance on only the youngest children (3-4 years) who met criterion across conditions, no significant results were found for non-misleading forced choice source, combination misleading source or combination misleading detail questions: $t(12) = 1.54$, p = *n.s.*, $t(12) = .49$, p = *n.s.*, $t(12) = .18$, p = *n.s.*. However, when children who have not met criterion are removed and entered into planned independent t tests, the small sample size (only approx. 8 per cell) makes it difficult to find significant effects. Means for the criterion group were consistently higher than those of the control groups across all question types: (*Ms* = .68, .25, .40) for criterion group and (*Ms* = .47, .18, .37) for control group on the three question types respectively.

Table 5: Means (and standard deviations) of target source question accuracy across age groups and conditions for participants who did not meet criterion in any condition, N= 41).

		Question type	
Age x Condition x Question Type	Non-misleading forced choice source	Combination misleading-source	Combination misleading-detail
3- to 4-year-olds			
Control	.42 (.24)	.22 (.31)	.37 (.36)
Criterion-training	.68 (.24)	.25 (.18)	.40 (.33)
7- to 8-year-olds			
Control	.67 (.20)	.30 (.15)	.68 (.21)
Criterion-training	.64 (.14)	.50 (.13)	.62 (.17)

Discussion

In Study 1, the two hypotheses were generally supported. First, training of any type (set amount or criterion) improved the older children's source accuracy. Second, training benefited the 3- to 4-year-olds only if they met criterion (four consecutive questions correct), and only when answering questions in the same format as that in training (i.e., forced-choice source questions). The data revealed developmental differences in children's ability to engage in source monitoring. The findings also resolve an outstanding issue – whether young children can benefit from source-monitoring training. Our results, taken together with the small number of training studies in this age group (e.g., Poole & Lindsay, 1995; Thierry & Spence, 2001), show that young children do have the cognitive capacity to benefit from source training but are not cognitively able to do so unless intensive training and feedback is given. Older children, on the other hand, could spontaneously engage in source monitoring after little instruction.

Developmental Differences

The developmental differences in source-monitoring training are consistent with the source-monitoring framework (Johnson et al., 1993). Although children as young as 5 can carry out some implicit source judgments automatically (e.g., Whitcombe and Robinson, 2000), more effortful and systematic source processing is not observed until later (Roebers, 2002). The televised and live sources in the current study was likely to be difficult because the sources contained similar perceptual information and thus automatic source monitoring (of which younger children are capable) would be less useful than more strategic source monitoring (of which older children are capable). While all of the 7- to 8-year-olds may have been capable of more systematic source decisions, only the ones who were encouraged to do so through training actually seemed to engage in such effortful processes (as indexed by source-monitoring scores). This raises the possibility that, for this age group at least, a prerequisite for accurate source monitoring in a difficult task is an awareness that more stringent evaluation of memories is needed. The 7- to 8-year-olds in the control group were probably capable of systematic source decisions, but unaware that more stringent evaluation was needed. Thus, it might be fruitful in the future to examine children's awareness of the difficulty of source-monitoring tasks at different ages.

It is encouraging that training benefited the 7- to 8-year-olds without any identified negative side-effects (i.e., the training groups were consistently more accurate than children in the control group on questions probing source, and were also more likely to correctly reject the suggestions in the misleading recognition questions). The training may have encouraged the 7- to 8-year-old children to "consider all options" before responding to source questions, rather than responding automatically on the basis of familiarity only. Reports based on familiarity-driven processes are common in research on children's memories of multiple sources (e.g., Powell et al., 1999; Ruffman, Rustin, Garnham, & Parkin, 2001). For example, in Powell et al.'s study, 3- to 8-year-old children experienced an event four times and were later interviewed about the final occurrence. Children rarely incorrectly claimed that they had seen a new detail that was not in any events, but were able to correctly identify 75% of the details that were in the events. Of these 75% of identified details, however, half of them were actually from the three occurrences prior to the final occurrence. Thus, children of this age appear to prefer familiarity-driven processing, but the results of the current study and others (e.g., Poole & Lindsay, 2001, 2002) suggest that it is relatively easy to encourage 7- to 8-year-olds to engage in more strategic source monitoring in lieu of such familiarity processing. Training effects for older children in the criterion condition were seen in responses to the misleading-source questions, while those in the set-training condition excelled in response to nonmisleading forced-choice source questions. It is possible that criterion training made children conservative in general resulting in superior performance on the misleading (possibly harder) questions, whereas set training improved source monitoring but restricted it to the questions most like those in training which also happened to be easier than misleading-source questions.

Question Difficulty

The training in the current study appeared to facilitate source memory when the children were asked very difficult questions. The misleading-source questions were particularly difficult for children (see Table 5), as they misled the source of an event that the children actually saw (e.g., “In the video, did someone pretend the banana was a telephone” - when this actually happened in real life game). The extremely low mean scores on these questions were likely a result of children answering “yes” based on the familiarity of the item (from recognition memory), and thus disregarding the source element of the question. Further, if children were not answering based on familiarity of recognition, we would expect that the misleading-detail questions would be similarly difficult for children (e.g., “In the real life game, was there orange water mixed with the crackers?” – when only blue water was mixed). However, the questions misleading the source of an actual event (or item) that happened were significantly more difficult for children to reject than those misleading the detail. This suggests that children had a recognition memory for that item or event being probed, and were using a familiarity strategy to judge ‘if’ it happened, and were ignoring the source element of the question. This finding has implications for the use of difficult, multi-clause questioning in forensic interviews (e.g., “Did he take pictures of you *in the bedroom*?”), as children may not fully attend to the ‘where’ component of these question types. The italics indicates the source element of the question which younger children may be inclined to ignore when answering. It should be noted though, that further study is needed on this finding, as this study was not designed specifically to answer this inquiry.

As is clearly shown in the results of Study 1, children found the misleading combination-source questions to be particularly difficult. The misleading-source and the misleading-detail questions were complex in that they contained multiple clauses. In general, children find questions with multiple clauses (i.e., the ‘combination’ questions) to be more difficult to answer than questions with a simpler structure (Walker & Warren, 1995). Responses to the combination questions suggest that, without training, children were responding on the basis of familiarity alone and not paying attention to the source component of the questions.

Responses to the combination-detail questions were clearly more accurate than responses to the combination-source questions. Recall that the combination-detail questions misled that a detail that was never present in either source occurred in one of the two sources, but the combination-source questions misled that a detail that was present in one source occurred in the other. Thus, responding on the basis of familiarity (saying 'yes' as soon as the detail is recognized) would result in good performance in response to the combination-detail questions, but poor performance in response to the combination-source questions, and this is exactly what happened when children were not trained to monitor source. When trained, however, responses to the combination-source questions were improved relative to a group of children who were not trained in source discrimination. Training benefits were not seen in responses to the combination-detail questions probably because a) source identification was not needed for these questions, and b) baseline performance on these questions was relatively good, thus there was less opportunity to observe a training effect. The finding that memory for content and memory for source are dissociated raises concerns for forensic interviews as combination-detail type questions are relatively common (e.g., Did he take off your clothes last time?). Presumably, a 'yes' response may be given if a detail is remembered (i.e., a person removing clothing), with no regard to the source component of the question (last time). Indeed, children are more suggestible when a suggestion is linked to a specific occurrence (e.g., "the last time?") rather than a series of occurrences ("Did that ever happen?"; Powell, Roberts, & Thomson, 2000).

In sum, the results of this study are encouraging in that training 7- to 8-year-olds in source monitoring transferred to a different task and resulted in more accurate reports. Source-monitoring training also benefited younger children aged 3- to 4-years but its effects were limited and context-specific. Specifically, training only helped when children had been given enough training to reach a specified criterion and benefits were limited to the exact format of questions on which they had been trained. Importantly, source-monitoring training did not hurt either age group: the means of children in the training groups were consistently higher than those of the control groups, even if they did not reach a level of statistical significance. The current results taken together with other research (e.g., Poole & Lindsay, 2001, 2002; Thierry et al., 2001; Thierry & Spence, 2002) clarify that a) older children can benefit from any kind of source-monitoring training, and b) some younger children may have the conceptual prerequisites to benefit under specific circumstances. The results provide a context for conceptualizing the developmental pathway of source monitoring in terms of the automaticity of strategic source processes and metacognitive awareness of source difficulty. Future research could focus on identifying what determines readiness to benefit from source-monitoring training and focus on flexible training protocols that adapt to children's current level of functioning.

Difficulties establishing criterion for source training

The results with the younger children identify one reason for the inconsistent results of source-monitoring training with this population. Studies providing children with a set amount of training report no training benefits (e.g., Poole & Lindsay, 2001, 2002), and we replicated this. Training benefits have been reported in other studies when 3- to 4-year-old children were trained until they reached a specified criterion (e.g., Thierry & Spence, 2002). In the current research, we found that training reduced source confusions in this age group *only* when they were trained until they reached criterion (even though the training was less intensive than in Thierry & Spence's study). Thus, it appears that young children need more than a simple orientation to source but rather require intensive accuracy feedback and practice until they have achieved competence in a given source-monitoring task (in contrast to the older children). Although the criterion was arbitrary, it probably indexed children's competence in mastering the source-monitoring strategy. Unlike traditional 'strategies' used in memory studies, the strategy taught in SMT is a global use of more 'effortful processing'. Thus, it may have qualitative characteristics, and may slightly differ across participants. Nonetheless, this more scrutinizing examination of memories seems to make a difference for some children in the accuracy of their memory recall. Although there are other methodological differences between the small number of studies in this area, the current results show that young children's previous inability to benefit from source training may reflect performance deficits (whereby they need more training and instruction to monitor sources spontaneously) rather than underlying conceptual limitations (whereby no amount of training will reduce source confusions).

The younger children in this study only benefited from training when responding to forced-choice source questions. Perhaps these nonmisleading questions were easier to answer than the other question types or perhaps the benefit was because these questions were in the same format as those that they were trained on (and, hence, easier to apply the newly-acquired strategy, cf. Brown & Kane, 1988). The means of the older children who were trained were higher when responding to forced-choice source questions than misleading-detail questions, but responses were equally accurate to these two question types in the control group. In contrast, the younger children were consistently more accurate to the forced-choice questions than the other questions. Note that the young children in Thierry and Spence's (2002) study also showed training effects only on the question types used in training. This lack of transfer of training shows that while young children can be trained, they need intensive training and the benefits may be limited. Nonetheless, the results are encouraging, and the facilitative effects of training were not limited to these question types with the older children, and these effects are consistent with those in the strategy training literature (e.g., Bjorklund et al., 1997; Miller & Seier, 1994).

One difficulty in assigning "criterion" level is that the arbitrary criterion point chosen may be too high for younger children (3-4 years), and/or too low for older children (7-8 years), as there is significant developmental change during this period. Further, one 3-4 year old child might be challenged by the established criterion, whereas the next might find it easy. This occurred in the current study. Nearly all of the 7-8 year old children in Study 2 met criterion easily after four questions, whereas several of the younger children did not meet criterion after 20 questions. It was clear to us that questions that challenged older children, were seemingly too difficult for younger children (and vice versa). Thus, although a "criterion level" had been established, it was not sensitive to the individual differences of each child. As a result of this discovery, Study 2 examined a "levelled" criterion training system in which questions increased in difficulty if a child answered correctly, and got easier if a child answered incorrectly (much like the administration of the computer based GRE; Educational Testing Service, 2007). It was believed that such a system would make questions more likely to fall within a child's "zone of proximal development" (Vygotsky, 1978).

Thierry and Spence (2002) has been the only published SMT study to introduce some element of "criterion" into training. They suggest that this was a possible factor explaining the null effects they found in training younger children in the Thierry et al., (2001) study. If children are not at least within the utilization deficiency phase, they cannot benefit from SMT, as they cannot even reproduce (with success), the strategy. Thierry and Spence (2002) indirectly suggest that facilitating the transition from "utilization deficiency" to "production deficiency" requires challenging children's source monitoring to some criterion level and allowing for sufficient practice. This is supported by Brown and Kane (1988) who suggested that young children need intensive and supportive training to benefit from strategy usage. Establishing a "minimum level" of training, might (as Thierry & Spence, 2002 suggest) ensure that children have received the minimum amount of training they require to use a strategy spontaneously (in a target interview). Although Study 2 will examine SMT using a criterion level with children, it is difficult to know exactly *where* one might establish a criterion for 3-8 year old children. For instance, it is plausible to suggest that a three year old child might require a different level of "criterion" and support to benefit from SMT, than does an eight year old (who might spontaneously use a strategy even with minimal source training). To ensure that children are given training which is "supportive" and "intensive" as Brown and Kane

(1988) suggest, a system must be developed that matches and builds upon a child's current level of cognitive functioning and strategy usage. If a criterion system is improperly gauged, it may be too difficult for younger children and too easy for older children (a problem not yet addressed in the literature even though several studies have used the same SMT with various ages of children). It is plausible to suggest that an ideal training system would "scaffold" a child's current ability to monitor the sources of their memories, and build upon this in a gradual and individually flexible way. Scaffolding involves working from (but not above) the child's current level of ability, and within their "zone of proximal development" (Vygotsky, 1978). In short, an appropriately scaffolded training procedure would individually match a child's abilities, but not exceed the upper bounds of their current abilities. Study 2 examined the development of a SMT system which will administer questions of varying difficulty in a "levelled" manner to children to allow individualized training. This system was intended to maximize the challenge for a child, but still stay within their cognitive scaffold of current abilities (rather than being too difficult or too easy). Although 3-4 year old children may be able to *do* a SMT task in training (utilization deficient), it is believed that more supportive and intensive training (Brown & Kane, 1988) might assist these children in making the transition to using these strategies spontaneously on a subsequent task (production deficient).

Without establishing appropriate criterion levels in SMT, researchers may "under train" children and not facilitate the transfer from utilization deficiency to production deficiency (as Thierry & Spence, 2002 suggested was the case in Thierry et al., 2001). Although it seems less likely to "over-train" older children with SMT (as older children seem to benefit from even subtle training), this is an unexplored domain, and it is plausible to suggest that "over-training" might fatigue or discourage children (from repeatedly answering incorrectly) as suggested in Bjorklund et al., (1997). It may also be that younger children do not benefit from "over training", as they do not understand the strategy being taught and thus, cannot use it.

Study 2

All previous research on SMT that has examined multiple age groups has used the same SMT procedures for all ages of children. Results from Study 1 examined if SMT is effective in increasing source accuracy for both younger and older children, when the same SMT is used for both age groups. It is possible however, that SMT used with older children may not be appropriate for use on younger children. It is plausible to suggest that due to the huge gains in cognitive development and executive function between 3 and 8 years, that training questions which challenge a 3 year old might not be appropriate for an 8 year old (and vice versa). The purpose of Study 2 was to develop and test a multi-Tier source monitoring training system that might be flexible to individual cognitive differences. Such a system may challenge children at their "optimal" cognitive level regardless of age.

Study 2 piloted and tested a multi-Tier SMT system which operates on a 'sliding scale' of criterion. The exact question given to each child was determined by the accuracy they had shown on previous questions. This attempted to ensure that children were trained to their 'optimum' cognitive abilities, but not significantly under or over this level. By the end of administration, it was assumed that they were within their 'optimum range' of correctly answering questions. Such a training system is not unlike the computer-based Graduate Record Examinations, which uses a levelled system of questions, with the difficulty of questions determined immediately as the student is writing (Educational Testing Service, 2007).

Materials

Pre-developed scripts and question sheets were used for all sessions (see Appendix A for full list). Video cameras and audio recorders were used to record all sessions. Materials for the lessons included various materials that were purchased at educational stores and dollar stores fitting the 'theme' of the script (e.g., nature lesson, frog lesson, digestion lesson).

Participants

One hundred and thirty six children in three age categories were used to examine the effectiveness of the developed multi-Tier SMT procedure. Forty-three children aged 3-4 (M= 3 years, 10 months, ranging from 36-55 months) , forty children aged 5-6 (M=6 years, 1 month, ranging from 61-83 months) and forty-eight children aged 7-8 (M= 7 years, 11 months, ranging from 84-106 months) were given the same target event and video as in Study 1, and subsequently assigned to criterion training, set training, or recognition training (control) groups. Written parental consent and verbal child assent were received prior to any child participation. Children were recruited from the Laurier Families Database (a list of families who have expressed interest in participating in research), as well as one local elementary school and one local daycare.

Procedures

Session 1 – Target Event

During session 1, children watched one of two (order counterbalanced) versions of the same "All about Digestion" lesson and video (order counterbalanced) used in

Study 1. To maintain consistency across studies, the same scripts, videos, props and live actor were used in these lessons. As in the Study 1, children watched the lessons in small groups (3-4 children) and were then returned to class. Children were randomly assigned to a "yoke group" group prior to Session 2. Each match group consisted of 3 similar aged children: one was randomly assigned to the criterion training group, one to the set training group, and one to the control group. Each child was asked the same questions (training and target questions) as their yoke in the same order. This procedure ensured that even though the set and control groups were not given "individualized training", they were given the same number and type of training questions as the yoke in their criterion group (thus removing a confound of individualized training procedures). Once children were yoked for age and training, the criterion child in a yoke group was run first for Session 2 (as the criterion's training was individualized, it determined the training for the set and control yokes to be run next). The subsequent questions asked of the yoked set and control children were based on those of the previously run criterion child (see below). Without matching children (based on age) across training groups, the criterion children may have received significantly "more or less training" than other groups, as there is no predetermined number of questions for any one child (children could have received a range of 9-25 training questions). Such a difference would have confounded training results across groups. The yoking procedure ensured that all groups received on average, the same number, order and type of training questions. Both the target events and interviews were videotaped.

Session 2 – Training and Target Interview

There was a 2-3 day delay between session 1 and session 2. During session 2, children were approached by a different research assistant (similar to Study 1). Children were asked if they wanted to learn about nature by watching a video and live lesson. Once assent had been given, children were taken individually to watch the nature lessons and video developed and piloted for Study 2. After the video and live lesson, the props used in the lesson were covered with a blanket and training began.

Study 2 developed SMT questions for children across four Tiers of difficulty. The development of these Tiers was based on previous research as well as an extension of the results of Study 1. Their structure was as follows:

Tier 1 questions – These questions were deemed to be the 'easiest' for children to answer. Although undoubtedly each individual question will vary in difficulty based on the memory of the individual target item it asks about, the format and structure of Tier 1 items overall, should be easier to answer correctly. Tier 1 items asked children about non-misleading items which occurred in one source (TV or game) with no similar instantiation in the other source. (e.g., "Did you see a tornado in a pop bottle?" – When this happened in the game only). If a participant answered "yes" to the recognition question in all tiers of training, a follow-up source question was subsequently asked (e.g., "Was that in the real life game or the video" (order counterbalanced). Note that participants in the control condition did not receive the source follow-up question, as this could inadvertently provide cues to monitor the sources of their memories more closely and thus, provide them with source training.

Tier 2 questions – Tier 2 questions only differed from Tier 1 questions in that they asked about items that did not occur or were not presented (misleading items). When answering these questions, children must 'correctly reject' these novel and misleading items (e.g., "Did you learn that leopards live in high, cold mountains" – When no mention of leopard took place in either source). This "correct rejection" element of the Tier 2 questions should have been more difficult for children, as it requires more effortful strategies to reject a misleading suggestion (Poole & Lindsay, 2002; Thierry & Spence, 2002). Misleading questions were introduced as a more difficult question type, as it is more difficulty fro children to correctly reject a misleading question than to accept a non-misleading question. Secondly, having a mix of misleading and non-misleading questions across tiers reduces the chance that children will succeed in progression by simply saying "yes" during training (i.e., a yes bias).

Tier 3 questions – Tier 3 questions were non-misleading questions that were asked when a plausible instantiation took place in the 'wrong' source. That is, the correct recognition answer is 'yes' but participants must then choose between two plausible sources in order to answer the source follow-up (e.g., Did someone use crayons to draw? – When someone used crayons in the game, but paint in the video). It was predicted that these 'plausible instantiation' questions would be more difficult for children as they presented very difficult source discriminations for them.

Tier 4 questions – Tier 4 questions were misleading version of Tier 3 questions in that they presented an option which was a plausible (based on the 'script' or 'gist' memory) but differed in that the option presented was not the correct plausible instantiation (and thus is misleading). For example, children might have been asked "Did you see a green leaf?". When they actually saw a red leaf in one source (game) and a yellow leaf in another source (TV). These questions were predicted to be very difficult for children to correctly reject, as they would presumably have two source memories of learning about leaf colour – thus making 'green leaf' a plausible (or 'gist consistent' option). The use of these question types in this study was somewhat exploratory, as they have never before been examined in source monitoring training research. However, they are forensically relevant question types, as it is easier to imagine 'plausible instantiations' being used in eyewitness questioning, and these have been found to be very difficult for children (Roberts & Powell, 2006).

The question types used in Tiers 1 and 3 were comparable to those used in training for Study 1. These questions present non-misleading questions with (Tier 3) and without (Tier 1) plausible instantiations given in the other source. Questions across all Tiers relate to the target question types used in the digestion interview, as questions with both instantiations and no instantiations were used as target questions.

Criterion Training

Children in the criterion training group began receiving training questions at Tier 1 (see Appendix A for complete training questions, scripts and target questions). Training was stopped if a) a child received a total of 25 questions or b) if a child received all questions at one Tier (maximum 9 at any one Tier). Children progressed up and down the Tiers of questions based on the following criteria: children moved "up" one Tier if they answer two consecutive questions correct, and children moved "down" one Tier if they answer two consecutive questions incorrect. If an "I don't know" response was given, it did not affect the progression of Tiers, and was not scored as correct or incorrect. This decision was made as the instructions given to children encouraged them to say "I don't know" if they were unsure of a question (similar to a forensic interview). A question was scored "correct" if both the recognition and follow-up source components of the question were accurate (see Appendix A for scoring sheet). If either the recognition or source component of the training question was incorrect, that question was scored as "incorrect". If a child answered a recognition question correctly, they were reinforced with the following statement: "That's right, you did learn about (target item)". If a child answered a recognition question incorrectly, they were corrected with the following statement "That's a good try, but actually, you did learn about (target item)". After the recognition training component of the question, children in the criterion and set training group received a follow-up source question. They received accuracy feedback in a similar manner to the recognition component of the question (e.g., "That's right, you did see (target item) in the video"). If children had answered a total of 5 incorrect questions at Tier 1, they received an additional training statement (see Appendix A for training script). These additional training statements were intended to help children who were not doing well at Tier 1 (as there is no lower Tier for easier questions). Children in the criterion training group progressed up and down Tiers until one of the aforementioned rules of termination was met.

Set Training

Children in the set training group received the same recognition and source training questions as their "yoke" did in the criterion group (in the same order). This ensured the exact same questions (and same order) were given to the age-yoked children in these groups. Individualized accuracy feedback on recognition and source questions was still given throughout their training similar to the criterion group. Training stopped whenever training stopped for their criterion yoke. The only difference between the set training and criterion training groups was that the criterion group received an 'individualized, set of questions' and that the set group received pre-determined questions based on their yoke. So, children in the set training group received the same training questions as their "criterion yoke", but the feedback given (e.g., that's right") was based on the accuracy of their response (rather than that of their yoke). Thus, although both training groups receive accuracy feedback, only the criterion group's training will be determined by their accuracy in training (thus, it is scaffolded to their ability). This ensured that both training groups received correct accuracy feedback of their answers, and the same number of training questions.

Recognition Training (Control)

Children in the recognition training group received the same recognition training as their yokes had received in the set and criterion groups. They were not however, given any source follow-up questions in training. The only difference between the recognition training and set training groups was that the source follow-up questions were removed. These children however, received the same recognition training (same questions, same order) as their criterion and set yokes. Training stopped whenever training stopped for their set and criterion yokes.

Target Interview

After training had finished, children received the same transition script used in Study 1 from the interviewer explaining that they would like to know "what happened the other day with [experimenter]" (see Appendix A for script). After the transition script, children were asked 32 target questions from one of two counterbalanced lists (order counterbalanced) without any accuracy feedback (as the interviewer was blind to the correct answers). These target questions asked about details in the "All about Digestion" lessons, and were identical to those used in Study 1 (see Appendix A for list). Children were given encouragement "You're doing great", after every fifth target question. Session 2 was video and audio taped. After the 32 target questions were administered, the child was thanked and returned to their class. Children later received small treat bags as compensation for their participation and the participating schools received a small financial donation.

Although the training session differed across Studies 1 and 2 for experimental purposes, it is important to note that all other aspects of the methodology were held constant. For instance, the target events, target videos, items used, delay time and eventual target questions were consistent across studies. This consistency allowed for some exploratory cross-study comparison of training programs.

Results

The results are comprised of descriptive information of how children progressed in training, followed by analyses of target questions, and finally, a post-hoc comparison of Studies 1 and 2.

Training Results

As children were yoked across conditions, they received the same training questions (presented in the same order) as their yokes. Thus, no analyses for condition were necessary as each condition group were given the exact same training questions in the same order.

Not surprisingly, older children progressed to higher Tiers than younger children. The breakdown of "how high" each age group went in training were as follows:

Table 6. Cross-tabulation of highest Tier reached in training based on age group.

	3-4 years	5-6 years	7-8 years	Total
Tier 1	9	0	0	9
Tier 2	25	6	0	31
Tier 3	9	6	0	15
Tier 4	0	33	48	81
Total	43	45	48	136

As shown in Table 5, none of the youngest children (3-4 years) reached Tier 4, and all of the oldest children (7-8 years) reached Tier 4. As shown in Table 7 below, a similar pattern of Tier succession is seen when examined based on the actual Tier the child stopped at in training. One difference however, is that although 36 children in the youngest age category did move to Tiers 2 or 3 at some point in training, only 3 of these actually stopped training at these levels. The majority of the youngest children ended up back at Tier 1 (as errors at higher Tiers eventually moved them back down).

Table 7. Cross-tabulation of what Tier children stopped at in training based on age group.

	3-4 years	5-6 years	7-8 years	Total
Tier 1	40	6	0	46
Tier 2	0	3	0	3
Tier 3	3	4	0	7
Tier 4	0	32	48	80
Total	43	45	48	136

Target Questions

Responses to target non-misleading recognition and misleading recognition questions were analyzed together in a 3 (age) X 3 (condition) X 2 (question type) MANOVA (see Table 7 for means and standard deviations). Overall, there was an effect of age $F(2, 136) = 9.18$, $p < .05$, but not condition $F(2, 136) = .20$, *n.s.*. On non-misleading forced-choice recognition questions, a main effect was found for age with $F(2, 127) = 3.54$, $p < .05$ showing surprisingly, that younger children (3-4 years) were significantly more accurate than middle-aged children (5-6 years) on these non-misleading questions. As a result of these findings, data were checked for yes-biases. Three children in the youngest age group (3-4 years) were found to have said "yes" to all questions. After removing these children from the analyses, the aforementioned age effects were no longer significant with $F(2, 124) = 2.35$, *n.s.* On all other subsequent analyses in this study, removing these participants did not affect results, so these participants were included to increase power for analyses. See Table 8 for mean accuracy scores (and standard deviations) in response to target questions.

Table 8. Means (and standard deviations) of target recognition question accuracy across age groups and conditions in Study 2 (*N*=136).

	Non-misleading forced-choice recognition	Misleading forced-choice recognition
3-4 years		
Control	.74 (.19)	.39 (.28)
Set	.68 (.28)	.41 (.28)
Criterion	.72 (.21)	.37 (.31)
Total	.72 (.22)	.39 (.29)
5-6 years		
Control	.57 (.20)	.62 (.28)
Set	.57 (.25)	.56 (.28)
Criterion	.65 (.26)	.57 (.26)
Total	.60 (.24)	.58 (.27)
7-8 years		
Control	.63 (.14)	.74 (.20)
Set	.68 (.18)	.67 (.18)
Criterion	.63 (.21)	.67 (.17)
Total	.64 (.18)	.69 (.18)

On the misleading forced-choice recognition questions, a clear and significant age effect was found with $F(2, 127) = 16.82, p < .01$. This effect showed that the youngest age group was significantly lower than the other two older groups.

There was no significant condition effect found on these question types with $F(2, 127) = .20$, *n.s.* for non-misleading recognition questions, and $F(2, 127) = .44$, *n.s.* for misleading recognition questions. No significant interaction effects were found for either question type. When recognition scores were entered as a covariate, none of the training effects across any analyses changed.

For source questions, a 3 (Age) x 3 (Condition) x 3 (Question type) MANOVA was conducted. Numbers in these questions were fewer, as participants were only asked source follow-up questions if they responded "yes" to the previous recognition question (i.e., if a participant says "no I don't remember that", it would not be a valid source follow-up to ask “was that in the video or real-life game”). Included in this analysis were: non-misleading source follow-up questions, combination misleading source questions, and combination misleading detail questions (see Table 9 for means and standard deviations). For non-misleading source follow-up questions, results showed a significant age effect with $F(2, 125) = 21.40, p < .01$. This age effect showed that 7-8 year old children were significantly more accurate on these question types than both of the other age groups. For combination misleading source questions, a significant age effect was also found with $F(2, 125) = 7.12, p < .01$. This age effect showed that both 7-8 year old and 5-6 year old children were significantly more accurate than 3-4 year old children on these questions. A similar age effect was also found on combination misleading detail questions with $F(2, 125) = 35.05, p < .05$. Age effects on these question types were identical as the aforementioned combination questions, with both 7-8 and 5-6 year old children showing significantly better accuracy than 3-4 year old children.

Table 9. Means (and standard deviations) of target source question accuracy across age groups and conditions in Study 2 (*N*=134).

	Non-misleading forced-choice source	Combination misleading source	Combination misleading detail
3-4 years			
Control	.50 (.16)	.22 (.17)	.25 (.19)
Set	.55 (.22)	.25 (.28)	.40 (.26)
Criterion	.49 (.17)	.26 (.26)	.34 (.30)
Total	.51 (.18)	.24 (.24)	.34 (.26)
5-6 years			
Control	.49 (.30)	.38 (.21)	.62 (.26)
Set	.57 (.24)	.36 (.23)	.43 (.20)
Criterion	.59 (.15)	.42 (.19)	.56 (.22)
Total	.55 (.23)	.39 (.21)	.54 (.24)
7-8 years			
Control	.79 (.18)	.39 (.13)	.80 (.14)
Set	.73 (.17)	.40 (.20)	.71 (.17)
Criterion	.78 (.21)	.37 (.16)	.64 (.17)
Total	.77 (.18)	.39 (.16)	.71 (.17)

No significant condition effects were found on these question types for: non-misleading source follow-up questions $F(2, 125) = .24$, *n.s.*, combination misleading source questions $F(2, 125) = .09$, *n.s.*, or combination misleading detail questions $F(2, 125) = .52$, *n.s.*

One significant Age x Condition interaction was found on combination misleading detail questions $F(4, 125) = 2.93, p < .05$. Follow-up univariate analyses were conducted to explore this interaction. Surprisingly, it was found that this interaction stemmed from the control condition having significantly higher scores than the criterion condition in the oldest age group (7-8 years) with $F(2, 45) = 3.83, p < .05$. A similar marginal effect was found in the 5-6 year old age group on this question type with $F(2, 42) = 2.78, p = .07$. Interestingly, on this question type, the control condition had the highest accuracy in 7-8 and 5-6 year old children.

Comparison across studies 1 and 2

As the means for all groups in Study 2 seemed high in comparison to those in Study 1, it was suspected that perhaps participants in Study 2 outperformed those in Study 1 (see Table 10 below). This would be a surprising finding, as no between-group differences were found between training and non-training groups in Study 2.

Although training was different across studies 1 and 2, the target details, target questions, and delay times were identical (not including the 7-10 day delay condition in Study 1). Thus, some comparisons were possible across studies for exploratory purposes. A 2 (Age) x 2 (Study 1 or Study 3) x 3 (Condition) MANOVA was conducted to examine trends across studies. As shown in Table 10, very few differences were found in 3-4 year old accuracy on the three source target questions. However, one key interaction was found. A significant Age x Study x Condition interaction was found for non-misleading source questions with $F(2, 160) = 3.31, p < .05$. Univariate ANOVA's were conducted to explore this interaction, which found that the criterion children in Study 1 outperformed their Study 2 counterparts on non-misleading forced-choice source questions with $F(1, 27) = 4.24, p < .05$. In general, training effects (or lack thereof), were consistent across studies in the youngest age group.

Table 10. Accuracy means (and standard deviations) for 7-8 and 3-4 year old children in Study 1 and Study across conditions (2-3 day delay, *N*= 94 for Study 1 and 91 for Study 2).

Age x Condition x Study	Question type Non-misleading forced choice source	Combination misleading-source	Combination misleading-detail
3- to 4-year-olds			
Control (Study 1)	.61 (.33)	.26 (.25)	.39 (.29)
Control (Study 2)	.50 (.16)	.22 (.17)	.25 (.19)
Set-training (Study 1)	.54 (.26)	.24 (.25)	.38 (.29)
Set-training (Study 2)	.55 (.22)	.25 (.28)	.40 (.26)
Criterion-training (Study 1)	.64 (.22)	.24 (.18)	.36 (.31)
Criterion-training (Study 2)	.49 (.17)	.26 (.26)	.34 (.26)
7- to 8-year-olds			
Control (Study 1)	.65 (.21)	.31 (.17)	.65 (.24)
Control (Study 2)	.79 (.18)	.39 (.13)	.80 (.14)
Set-training (Study 1)	.79 (.16)	.33 (.16)	.68 (.20)
Set-training (Study 2)	.73 (17)	.40 (.20)	.71 (.17)
Criterion-training (Study 1)	.64 (.14)	.50 (.13)	.62 (.17)
Criterion-training (Study 2)	.78 (.21)	.37 (.16)	.64 (.17)

In the oldest age group, a consistent trend did appear when accuracy was examined across studies. In this age group, Study 2 children in the control condition were significantly more accurate on: non-misleading forced choice source questions $F(1, 33) = 4.80, p < .05$, combination misleading source questions (marginal) $F(1, 33) = 2.29, p = .14$, and combination misleading detail questions with $F(1, 33) = 4.70, p < .05$. Thus, consistently across studies, Study 2 older children in the control condition were more accurate than their Study 1 counterparts. This suggests the possibility that the control children were inadvertently trained in source monitoring.

In the criterion condition for older children, the results were somewhat mixed when examined across studies. For example, although Study 2 criterion children outperformed their Study 1 counterparts on non-misleading forced-choice source questions with $F(1, 30) = 4.87, p < .05$, this outcome was reversed on combination misleading source questions; with Study 1 criterion children outperforming Study 2 children $F(1, 30) = 6.41, p < .05$. Across the set conditions, and all other cells, no other significant differences were found between Study 1 and 2 accuracy scores.

Discussion

In Study 2, the SMT was changed to support more individualized and 'scaffolded' training for the children. Our predictions that this may help facilitate source accuracy in older and younger children (exploratory hypotheses) were not supported using our multi-tier training system, as no significant training effects were found. However, exploratory examination of the data across Studies 1 and 2 revealed some encouraging findings, in that the means of Study 2 training participants were seemingly higher than would be expected. Although no clear effects arose from the study 2 'trained' versus 'untrained' groups, there was some evidence that older children may have shown an improvement from the Multi-Tier SMT procedure.

At first blush, it would seem that the training procedures did not provide assistance to children in any age group in Study 2, as few significant differences were identified across training groups. However, the possibility exists, that *all* groups received inadvertent training during the training session –even those in the control condition. Comparison of the accuracy scores across studies 1 and 2 somewhat supports this notion.

How might our control group have been trained in Study 2 (in addition to the pre-established training groups)? It is possible that the lengthier and more intensive training provided in Study 2 afforded a benefit to all groups; even those that were not given source follow-up questions. Keep in mind that even the control training group received on average, more questions than all groups in Study 1 (as the number of questions they received was determined by their criterion yoke). Thus, it may be an issue of both quality and quantity of questions given in training which might help to determine if a SMT procedure will be effective. As Brown and Kane (1988) suggest, training must give adequate support and practice for a person to benefit from it. It is possible then, that our shorter (and less intensive) training in Study 1 was sufficient for older children (7-8 years), but was simply not explicit or intensive enough for younger children. However, in Study 2, children were given more difficult questions in training (more difficult instantiations), as well as more questions to practice (and be given feedback) – regardless of their group assignment. As this multi-Tier training procedure has never before been used, and as very difficult 'instantiation' training has never before been tested in a SMT study, it is still somewhat unknown how these questions will affect a SMT procedure.

Nonetheless, these question types have the potential to arise in forensic interviewing, and thus, need exploration.

Comparison of Study 1 and 2

As the target events, questions and delay times were identical across studies, some exploratory study comparisons were conducted. It should be noted however, that results of these analyses cannot be considered definitive, as the training procedure was different in Study 2. When we compared the accuracy scores (across identical events, question types, and delays) in studies 1 and 2, we found evidence that perhaps older participants in the training groups benefited from our multi-Tier procedure. The results presented above suggest that the mean scores for children in study 2 were in general, significantly higher than those children we deemed as 'trained' in study 1. If mean scores for children are higher than a 'trained' group (using the same delay and target items), it is possible that somehow, all groups received training – thus washing out any 'training group' effects in study 2.

Although yoking participants in Study 2 controlled the confound of 'training intensity' (i.e., the number of questions asked and the order these are given), this may have also inadvertently ensured that all children were trained. If the level of training is a good proxy for SMT, Study 2 would not have captured this, as all groups were controlled in this variable. One issue that arises with comparing a criterion and set level of training (as Thierry & Spence, 2002 and Study 1) is that as children reach criterion, they are inadvertently "trained more". Although this makes good theoretical sense from a training perspective (as those children seem to need more training), it confounds the notion of 'quality' versus 'quantity' of training. In Study 2, our research controlled for 'quantity', but adjusted the quality of training (i.e., making it individualized and levelled). In doing this, we arrived at null training effects across groups. However, when accuracy means were compared with participants in Study 1 (who had the same events, delays and target questions), evidence suggested that our Study 2 participants outperformed our Study 1 participants (who were trained). This suggests the possibility that giving 'more' training, may be the proxy that assists younger children. However, our cross-study comparison was a post-hoc analyses, and needs further exploration, as the studies were not designed to answer this question specifically.

Although it may be that training might better be determined by 'more' questions in training, it should also noted that the qualitative characteristics of our questions were quite different in Study 2 at the Tier 2 and 4 level. At these levels, participants were given misleading questions (with and without instantiations), something they were not given in Study 1. Thus, at these levels, the 'depth' of training changed also. However, looking at Table 5 in the results section, it is apparent that only 9 out of 43 of our youngest children reached Tier 3, and none of them reached Tier 4. Similarly (as shown in Table 6), 40 out of 43 of these participants eventually stopped at Tier 1. Thus, the training for 3-4 year olds in Study 2 may have been more intensive, but the qualitative characteristics of the questions they were asked were not substantially different. In short, the types of questions asked at Tiers 1 and 3 were comparable to those that were used in training in Study 1.

General Discussion

The notion of whether children have the ability to benefit from source training has been a topic of discussion in recent years amongst source monitoring researchers, as such a system (if found to be reliable) would benefit forensic interviewers. Although some researchers have found that certain SMT programs are effective with younger children (e.g., Thierry & Spence, 2002), others have not replicated these findings (e.g., Poole & Lindsay, 2002). Thus, some debate still exists as to whether younger children can benefit from source training, and if so, under what conditions. As there are many differences between these aforementioned studies examining source training in children, it is difficult to compare them. The current study attempted to extract one variable (the notion of criterion-based and individualized training), and test this across several age groups of children which are commonly used in SMT studies.

Taken together, the purpose of these studies was to examine the effects of two different types of source monitoring training on different aged children. More specifically, we were interested in exploring if differing 'levels' of training had any effect on source accuracy of different question types.

In general, the results were encouraging, in that Study 1 showed consistent benefits of SMT for older children with any amount of training, as well as some benefit of SMT in younger children when they were trained to a criterion level. Initially, our analyses for Study 2 indicated that no significant effects of training were found, as our training groups did not differ significantly from the control groups (in all ages). This was perplexing, as training effects in older children (7-8 years) were seemingly easy to induce; as we had achieved this by only asking four questions in Study 1. As we had used identical target events and target questions across studies, an exploratory comparison of training programs was possible, and in general found no significant benefit of the training used in Study 1, over that used in Study 2. Thus, there is some speculative and preliminary evidence to support the facilitative effect of the training used in Study 2. In some senses, this training may be more efficient that that used in Study 1, as similar accuracy means were achieved without giving any accuracy feedback (in the control group); as was the case in Study 1. However, strong conclusions of training effects in Study 2 cannot be drawn, as no significant cross-condition differences were discovered.

Comparing Theoretical Explanations

Source monitoring theory suggests that when we make decisions about the origin of a memory (i.e., the source), we are making attributions based on the amount of effortful processing we have dedicated to that attribution (Johnson et al., 1993). That is, the greater effort we place in making that attribution, the more accurate our source monitoring should be. Source monitoring training is merely a means to induce more effort and/or attention in making the attribution. It has been suggested through previous research that very little prompting is required to effectively 'cue' a child into making more effortful attributions about source (Poole & Lindsay, 2002; Thierry et al., 2001; Thierry & Spence, 2002). Although some debate still exist as to the age at which children can benefit from this cueing, it is clear that when children can benefit, very little cueing is necessary to achieve results with older children. It is possible this was the case in Study 2, although this is merely a speculative notion based on exploratory analyses. It is possible that the significant increase in the amount of questions asked in Study 2 was enough to benefit all participants – even those who were not given source-follow-up questions. It is possible that if given enough practice with difficult questions, effortful processing can be facilitated, even without specifically asking source questions.

Keep in mind, that immediately after training, all participants were given source questions (and combination questions) about the target events. It may be that children generalized the effortful processing to the 'next step' of the interview when they were given source questions. Thus, children may have been given enough practice 'thinking hard' about questions, that when it came time to give the target interview, they were indeed trained.

Another interesting finding of these studies is the great difference in difficulty participants had with the different question types asked of them. Although 'question type' has not been heavily focused on in SMT studies, the current research suggests that it may be fruitful to continue this. As Walker and Waren (1995) suggest, questions which reference multiple clauses (i.e., combination questions), are extremely difficult for children, and are unfortunately regularly used in forensic interviews. Our research supports the idea that these questions are most difficult for children. When asked a question such as "In the video, did he drop a kidney on the floor?", the child may first attend to the detail component of the question (did he drop a kidney), and if confirmed in memory, ignore the pertaining source reference (in the video). The fact that the misleading source combination questions were significantly more difficult than the misleading detail questions suggests that this was the case (as we would expect these to be the same if children were not 'thrown off' by the source reference). Question types asking both a recognition component and a source component need further exploration, as they are asked within forensic or investigative interviews (e.g., Did he take pictures of you in the kitchen?).

The notion of combination questions affecting source monitoring accuracy might also be examined from a fuzzy trace perspective. Recall that fuzzy trace theory suggests that instead of a decision process (or attribution) about where a memory has come from, we 'pull' either a verbatim or gist memory (Brainerd & Reyna, 1990). Source accuracy is merely dependent on which of these traces are accessed at recall. Verbatim memories are memories of the detailed event, which retain 'source information'. Fuzzy trace theory suggests that verbatim memories decay quickly with time, and younger children are less likely to use them (Brainerd & Reyna, 1990; Brainerd & Gordon, 1994).

When answering the 'combination' questions, it is possible that the children in this study were using their gist memories to do so. For instance, if asked the question "Did he use square crackers in the real life game?", a child must attend to both the *source* (real life game) and the *detail* (square crackers) to make an educated response. Fuzzy trace theory would suggest that because this detail is "gist consistent" (i.e., it fits generally what took place), then children may be inclined to use their gist memories to answer the question. If this were the case, children would incorrectly answer "yes" (when in fact it was round crackers that were used). Thus, these combination questions may incorrectly prompt a child to use their gist memories, as the detail fits the general 'gist' of what took place. If however, we had asked "Did he use a toboggan in the real life game?", children would attempt to access their verbatim memories – as the notion of "toboggan" does not fit the gist of what took place. Thus, whether we ask about "gist consistent" or "gist inconsistent" details within these combination questions may very well define their difficulty, but this needs further testing. The fact that the 'combination misleading detail' questions were easier for participants than the 'combination misleading source' questions would support fuzzy trace theory in that participants may have been accessing verbatim memories (either containing the detail or not), and may have been making their judgement based on this. Reyna and Lloyd (1997) suggest that there can be interference effects between gist and verbatim memories. For example, if a combination question asked about a detail from one source, but 'suggests' it is in another, a child might accidentally 'pull' the incorrect verbatim memory, based on the gist consistency of it (i.e., there are several plausible verbatim memories which fit this "gist"). From this perspective, it seems logical that if participants were indeed accessing verbatim memories, that the 'combination misleading source' questions would be most difficult, as the detail asked about would be in the verbatim memory (not the case in the combination misleading detail questions).

The fact that these questions were more difficult also lends support to source monitoring theory. When a plausible instantiation or 'script consistent' item is offered, it would take more effortful processing to reject this. This is comparable to remembering if you ate cereal or toast for breakfast yesterday. This thought process requires more processing than deciding if you had cereal or hamburgers for breakfast (as one is not script consistent). Thus, it makes sense that accuracy scores for combination misleading-source questions would be most difficult.

Fuzzy trace theory might suggest that we can train participants to access certain verbatim memories (rather than gist), but it seems unlikely that this training would generalize to several other verbatim memories. For instance, in our training (as with other SMT studies), children were trained on non-target sources. Thus, in Study 2 they were cued to be more accurate about the nature lesson they had just learned about. According to fuzzy trace, we are training them to access the verbatim memory of that nature lesson. Why then would children generalize this training to benefit them on a subsequent (and different) verbatim memory access (the target interview). If children do indeed generalize *what they access*, it could be suggested that they are making more effortful processing in later decisions about their memories (i.e., the target interview). As fuzzy trace does not incorporate a 'processing' element, it seems unlikely that generalizing of any training would be effective. In order to benefit from training, a child must take what they have learned and apply it to a new setting (and new memories). If the fairly automatic 'memory pull' of fuzzy trace was indeed correct, this training of verbatim access would be limited to that particular verbatim memory (i.e., children would only benefit on the memory they are being trained on). Transfer of training requires intentional and effortful processing on the part of the participant.

The need for individualized source training

The notion of the more 'individualized' training used in Study 2 was taken from the Vygotskian notion of scaffolding (Vygotsky, 1978). It was predicated that due to the large individual differences that have been identified in children's general memory accuracy (Quas et al., 1997), the type of training provided for older children (7-8 years) might not be appropriate for younger children (3-4 years). Thus, a 'levelled system' of training was developed to address this. An ideal platform for training would have been work within the child's 'zone of proximal development', as a more difficult training system would be unusable for younger children and thus would be deemed within the child's 'zone of distal development' (Vygotsky, 1978). Although we did not see success in younger children using our multi-tier training program, there was some encouraging results when we compared means across studies. Finding a system that is flexible and sensitive enough to scaffold most young children, yet robust enough to reliable find training effects may be a difficult task. To further complicate this, it is entirely possible to confuse younger children with a complex or lengthy training system, as their attention and focus to the attribution process may be easily lost. This may have been the case in Study 2. Although we found some training effects in Study 1 (when children were trained to criterion), these results were not replicated in Study 2. It is possible that the simpler and more concise training provided in Study 1 was better suited for younger children.

Simply providing children with training, however, does not guarantee that children have acquired a given strategy. One index of whether a strategy has been trained is whether children meet an established criterion (Thierry & Spence, 2002). Thierry and Spence used a criterion and reported that all children met criterion relatively quickly (i.e., needing at most only a single additional trial to reach criterion). In contrast, the 3- and 4-year-olds in Poole and Lindsay's (2002) study did not show such competence. The samples in the current research appear to be more similar to Poole and Lindsay's sample in that the 3- to 4-year-olds in the criterion-training condition of the current study needed an average of eight trials to reach criterion. That is, the young children in our study needed more practice to reach criterion than the minimum number of trials (in this case, four). This shows that children in the criterion-training condition were not efficient at source monitoring before training. When we analysed data from only those children who met criterion, however, there were some training effects that were not present when all children in the criterion-training condition were analysed. It is possible, then, that the children who achieved criterion were simply more ready to benefit from training than children who did not meet criterion.

From a developmental perspective, it may be possible to categorize different levels of readiness to monitor source: some children may benefit from source-monitoring training but not yet be able to identify when systematic source monitoring is needed; at a more advanced level, older children and adults may be aware which source-monitoring decisions are difficult (e.g., when sources are similar, Lindsay et al., 1991; when delays are long, Poole & Lindsay, 2001), and put into place more stringent evaluation of their memories.

Given that most research on the development of source monitoring investigates these processes in relatively young children (aged 3 to 8), there are many unanswered questions about the development of source monitoring in older children. As with all source monitoring studies however, age effects are somewhat confounded by older children's increased familiarity with everyday objects and events which can facilitate their memory of these. Nonetheless, age effects are well established in source-monitoring research (see Roberts, 2000 for a review). It is possible that children in the training groups had better memory of the sources before training, but note that children were randomly assigned to condition and so this is unlikely. Further, training effects were observed when recognition scores were controlled.

The results speak to theoretical reasons why young children confuse the sources of information. Benefits in performance after training preschoolers in memory strategies are frequently unsuccessful. For example, Miller and Seier (1994) report that 90% of the studies they reviewed resulted in children failing to spontaneously use the strategy they had just been trained in (i.e., evidence of a "utilization deficiencies"). Utilization deficiencies can be identified in a number of ways, but is evident when there is no difference in performance between children who are strategic and non-strategic (Bjorklund, Miller, Coyle, & Slawinski, 1997). Thus, it could be argued that the results of the 3- to 4-year-olds (who did not meet criterion) in Study 1 reflect utilization deficiencies in using strategic source monitoring. That is, children who were strategic (set-training condition) performed no better on the target task than children who were non-strategic (control condition).

Prerequisites for source training

What might underlie "readiness" to benefit from source-monitoring training? Several researchers have suggested that an understanding of dual representation (that the same entity can be represented in two different ways (e.g., what something looks like and what something really is) is related to source monitoring (e.g., Giles et al., 2002; Poole & Lindsay, 2002; Roberts, 2000). An understanding of dual representation, and an understanding of mental states in general, has been linked indirectly to source monitoring in misinformation paradigms (e.g., Templeton & Wilcox, 2000; Welch-Ross et al., 1997). The idea is that resistance to misinformation is related to children's ability to reason that an event detail exists in reality (true representation), yet a suggestive interviewer can inaccurately describe the detail (false representation). Children who can reason about conflicting mental representations are less suggestible than those without such understanding. If resistance to misinformation involves monitoring two sources of information (the false suggestion and the true reality) as indicated by a substantial body of research (e.g., Ackil & Zaragoza, 1995; Poole & Lindsay, 2001, 2002; Powell &

Roberts, 2002; Thierry & Spence, 2002), then mental-state understanding may be related to source monitoring. Subsequently, children with mental-state understanding should be more likely to benefit from source-monitoring training than children who lack the conceptual prerequisite of an understanding of the mind. To date, however, there is no published research showing direct links between mental-state understanding and source monitoring (see Welch-Ross, 2000, for preliminary data).

Drummey and Newcombe (2002) found that children's source monitoring seems to develop substantially between the ages of 4-6 years, and that younger children <4 years had mistakes consistent with the type of 'source amnesia' found in patients with prefrontal lobe damage. Drummey and Newcombe (2002) suggest that this abrupt developmental improvement may be related to children's increased ability to bind the features of sources together (thus helping with source attributions made using these features).

Kovacs and Newcombe (2006) offer additional cues as to what may help define readiness to benefit from source training. Similar to the notion of dual representation, they suggest that perspective taking may play an important role in development of source monitoring. In their study, they found that 4-5 year olds children seem to benefit (in source monitoring) from taking the perspective of the to-be-remembered person during encoding.

Although age is currently the best proxy for source-monitoring skills (see Poole & Lindsay, 2001), many researchers have reported large individual differences in source-monitoring skills (Leichtman et al., 2000; Roberts & Blades, 1999). As source monitoring requires many metacognitive skills, it is yet unclear what developmental component/s of this skill might hinder young children's ability to benefit from source orientation (Johnson et al., 1993; Poole and Lindsay, 2002). For instance, differences in the use of retrieval strategies, theory of mind development, use of imagery, and lacking knowledge of metamemory have all been implicated as possible developmental hurdles to achieving accurate source monitoring at an early age (Ackil and Zaragoza, 1995; Johnson et al., 1993; Lindsay et al., 1991). Understanding the skill set needed for source monitoring may illuminate important developmental processes in children's readiness to benefit from training. Presumably if the necessary prerequisites of benefiting from source training are indeed developmental (which they seem to be), it seems likely that younger children might be able to benefit from this – but that the benefits might simply be less. Further, we would need to cater our source training to the developmental level of the child, in order to truly optimize their benefits of training. One cannot plausibly say for instance, that benefits from source training simply "appear" at age 7-8. Source monitoring develops continuously from about age 3, beginning with implicit skills and developing to explicit reflection on sources (e.g., Roberts & Blades, 1995; Whitcombe & Robinson, 2000). Thus, it may be fruitful to focus source training on individual need, rather than age specificity and arbitrary criterion assignment.

The purpose of the current studies was to help clarify the inconsistencies that exist in source monitoring training research. In some sense this was achieved, and in others, it was not. Previous research had shown inconsistent results as to whether younger children (>5 years) could be trained to better monitor the sources of their memories. Although clear support was given in Study 1 that children can (under the right conditions) benefit from a SMT procedure, Study 2 provided inconsistent support by asking "what are the conditions necessary for training to occur". Clearly, when simple training is used, older children can benefit from SMT. Poole and Lindsay (2002) and Thierry and Spence (2002), as well as the results from Study 1 have given support for this. It is also encouraging that these benefits, across all studies, come at no expense to the accuracy or quality of eyewitness reports. It remains somewhat unclear however, how easily younger children can benefit from a SMT procedure, and what conditions must be present for this to take place. Speculative examination of the results across both studies gave some support that perhaps children may benefit from an individualized training procedure (i.e., easier than that given to older children), and perhaps one that gives more practice with questions. However, this study was not designed to test these factors directly.

Future Directions & Limitations

It would be prudent to focus some future research on dissecting the issue of 'quality' versus 'quantity' of training, and how this might affect younger children's ability to benefit from this. Although in some senses, Study 2 did do this, it was not designed primarily for this examination, and thus, cannot be considered definitive. Study 2 would predict that more practice with an easier or individualized training program might suit the needs of younger children.

Results from Studies 1 and 2 also highlighted the difficultly in arbitrarily defining criterion levels in SMT studies. It was clear from the results of these studies, that younger children require vastly different training than do older children. Few developmental psychologists would be surprised by this finding, but it has not yet been explored in research. Establishing a 'criterion' is a somewhat dangerous endeavour, as we are as researchers defining that particular child "trained". However, there are many factors which affect the accuracy of eyewitness statements, even when children are "trained" (social pressures, question type asked, repeated questioning, repeated interviewing). Although the notion of 'criterion' fits the Brown and Kane (1998) requirements for transfer of training, it does not fit an applied model for use in children's memory studies. Significant individual differences exists, thus, many factors affect accuracy scores, and it is possible that researchers (or the general public) might be fooled into believing that 'training to criterion' is the panacea for increasing source monitoring accuracy. It is more prudent to acknowledge and respond to, the huge individual differences that exist not only across ages, but also across children of the same age. Defining SMT in terms of 'met criterion' or 'not met criterion', may be reducing a complex set of processes down to how children performed on a small set of questions. In some senses, this is how much of developmental psychology has reduced the notion of 'theory of mind'. If a child passes a simple test, they are often deemed as 'having theory of mind' (or not). This cognitive process is more complicated than simply 'having it' or not, as is the notion source monitoring training. We as researchers are at risk of reducing and simplifying complex cognitive operations to the point where they are defined by a simple 5 minute test, and should be cautious of this in the source monitoring field. Further, it should be emphasized that the use of forensic "tools" to increase children's eyewitness accuracy has a tumultuous past (e.g., the daycare scandals of the 1980's), and thus, we need to be cautious in adopting new techniques into the applied field without rigorous validation and testing. The field of source monitoring training is far too young at this point to be considered an applied tool, regardless of the results of published studies.

Further research should look to examine the issue of combination questions. It was interesting to discover how difficult these questions are for all ages of children, and yet concerning to think that they are used in forensic interviews (Walker & Warren, 1995). This finding has real forensic implications, as children in the current study seemed to disregard source-reference in these combination questions. Thus, when asking a child "Did he take your picture on Wednesday?", children might be inclined to answer based on the detail (take your picture), rather than the source reference (on Wednesday). Such a response might have serious legal implications if a defendant has a strong alibi on Wednesday, and thus could not have done it 'that day'. Results from Study 1 and 2 indicate that children might be inclined to say "yes" if this detail did indeed happen; even it was in another source (e.g., Wednesday).

Although the current studies do inform forensic interviewers in many ways, we are still not close to providing a valid and reliable training procedure for use with real forensic cases. The risks of prematurely using a procedure which has not been rigorously tested and validated were seen in the 1980's sexual abuse scandals. Many people served long prison sentences before being exonerated based on questionable forensic interview techniques. It is hoped that the current research can add to existing research in the field to amass a good knowledge of how children respond to questioning using different training procedures and forensic questioning techniques.

Appendix A: Digestion event and interview forms

Invitation to participate Script (Child – School/Daycare)

Hello (name), my name is (give name). I work at a university. A university is a really big school. As your parents explained to you before, I am interested in how kids learn things. Are you interested in learning about nature/digestion? If you want, we can go into the library and learn about nature by doing a live lesson and watching a video. Afterwards, we'll come back here to your classroom.

After the short lessons about nature/digestion, I'll ask you some questions about what you've learned. You can quit at any time you want. In a few weeks, we'll give a toy to your teacher for you, to thank you for your help.

Are you interested in learning about nature?

OK great, let's go learn some things about nature/digestion.

Digestion Script (Version 1 Game)

Today, we're going to play a game and watch a video about digestion. First, let's **play a game/watch** a video.

There are lots of different types of food. Some are healthy and good for us, and some are not. Here's a banana **(use the banana as a phone).** Look at me being silly using the banana as a phone. Here are some other fruit **(show two apples).** These are **two apples**. They're healthy for us. We also have a **box of juice** (show item). It's also healthy. With our fruit and drink, we can eat lots of other things. These are just a few types of food we can eat when we're hungry.

The first thing we need to do is get food into our stomach right? Do you think we could swallow these **square crackers** (show crackers) like this? Nooooo. That's why we need to chew up food before we swallow it…to help it fit in our stomach.
Look at these teeth (teeth model), these help break up our food into tiny pieces. **"Hello" (use teeth to say hello. "Hello Mr teeth" (talk back)**. See the teeth take a bite of food **(use teeth to take bite of orange).** Funny that the big teeth took a bite out of the **orange**.

We also make saliva in our mouth – it's made right here (show salivary glands on wall chart). Saliva helps make the tiny pieces of food slimy…so that they can slide down your throat better. Let's break up these **square** crackers with this **hammer** in the **coloured bowl**. Now let's pretend to add some saliva (mash up crackers with **hammer** in **coloured bowl** and add water **[blue water]** in a bowl). We'll use this **blue** water for saliva, and add the crackers to this **coloured bowl**.

Once food is in our stomach, strong acid help to dissolve it and break it down even more. This way, we can pick up the good stuff our body needs later on. Sometimes when our stomachs are dissolving food, it can produce gas. That's why we burp. Let's put a bit more water and a fizzer into the bowl to make gas (put water and small bit of alka seltzer into bowl). See how the gas is bubbling **(make gurgling noises).**

Now we mix up the crackers with this **metal** spoon.

Our tongues help us to **taste the food**.

"I'm so hungry, I could eat a **house**"

Our bodies send **more blood** to our stomachs when it's full to help us digest food.

Look at Danny the digester. This is where everything is happening (show on foam model) on Danny the Digester.

I'll just take out Danny's **stomach**...there, I took out Danny the Digester's **stomach.**

What's your favourite food Danny? (mouth with Danny) "My favourite food is **potato chips**".

(Ask kid) And what's your **favourite drink?**

Did you know that **cows** have four stomachs to help them digest food?

Some of the things we can't digest, we must filter out of our bodies. That's what the kidneys do – they filter bad stuff out of our bodies. These are your kidneys (show on chart). Let's add this **dirt** to this cup of water. See how it's all mixed together? Now when we pour it through this filter, all the **dirt** stays...and the water goes through (kids pour water and sand mix through filter). That's just like what our kidneys do – they filter out the bad stuff.

Now let's put on our digestion aprons. Here are all the organs we need to digest things. Let's try and place the organs in the right place our aprons. So let's put on the aprons here (assistance if needed putting aprons on).

First we'll start with the stomach...where does the stomach go (experimenter places **stomach** on apron). Now let's put on the squiggly intestines. (Assistance). Oops, I put the **squiggly intestines on upside down**. (Fixes it). Now let's put the kidneys on (same). Now let's put on the lungs so we can breathe.

These lungs look **leaves from a tree**. (Drop lung) "Opps, I dropped a lung". (pick up and put on child).

Now let's put our hands on our **stomachs** (assistance if needed). Good, you know where it is.

Well, it looks like you learned a lot from this. Thanks so much for helping.

Now we can **go back to class/watch a short video about digestion**.

Digestion Script (Version 1 Video)

Today, we're going to play a game and watch a video about digestion. First, let's **play a game/watch** a video.

There are lots of different types of food. Some are healthy and good for us, and some are not. Here's a banana **(use the banana as a guitar).** Look at me being silly using the banana as a **guitar.** Here are some other fruit **(show two apples).** These are **two apples**. They're healthy for us. We also have **some eggs and a piece of bread** (show items). They're also healthy. These are just a few types of food we can eat when we're hungry.

The first thing we need to do is get food into our stomach right? Do you think we could swallow these **round crackers** (show crackers) like this? Nooooo. That's why we need to chew up food before we swallow it…to help it fit in our stomach.
Look at these teeth (teeth model), these help break up our food into tiny pieces. **"Hello" (use teeth to say hello. "Hello Mr teeth" (talk back)**. See the teeth take a bite of food **(use teeth to take bite of banana).** Funny that the big teeth took a bite out of the **banana**.

We also make saliva in our mouth – it's made right here (show salivary glands on wall chart). Saliva helps make the tiny pieces of food slimy…so that they can slide down your throat better. Let's break up these **round** crackers with **my hand** on this plate. Now let's pretend to add some saliva (mash up crackers with **hand** and add water on plate) **[blue water]** in a bowl). We'll use this **blue** water for saliva, and add the crackers.

Now we mix up the crackers with this **metal** spoon.

(**Eat cracker**) "Ummm, Yummy cracker".

"I'm so hungry, I could eat a **car**"

Sometimes it hurts when **we get hungry**.

When our stomachs are **full**, they stretch out big like this stomach (blow air into empty stomach). This is **how full** our stomachs can be.

Our tongues help us to **swallow the food**.

Once food is in our stomach, strong acid help to dissolve it and break it down even more. This way, we can pick up the good stuff our body needs later on. Sometimes when our stomachs are dissolving food, it can produce gas. That's why we burp. Let's put a bit more water and a fizzer into this **pickle jar** to make gas (put water and small bit of alka seltzer into bowl). See how the gas is bubbling in the **pickle jar**.

Look at Danny the digester. This is where everything is happening (show on foam model) on Danny the Digester.

What's your favourite food Danny? (mouth with Danny) "My favourite food is **potato chips**".

Did you know that **Panda bears** only eat only type of food? It's a tree named **bamboo**.

Some of the things we can't digest, we must filter out of our bodies. That's what the kidneys do – they filter bad stuff out of our bodies. These are your kidneys (show on chart). Let's add this **dirt** to this cup of water. See how it's all mixed together? Now when we pour it through this filter, all the **dirt** stays…and the water goes through (kids pour water and sand mix through filter). That's just like what our kidneys do – they filter out the bad stuff.

Now let's put on our digestion aprons. Here are all the organs we need to digest things. Let's try and place the organs in the right place our aprons. So let's put on the aprons here (assistance if needed putting aprons on).

First we'll start with the stomach…where does the stomach go **(**experimenter places stomach on upside down on apron**)**. Oops, I put the **stomach on upside down**. (Fixes it). Now let's put on the squiggly intestines. (Assistance). Now let's put the kidneys on (same). Now let's put on the lungs so we can breathe.

(Drop kidney) "Opps, I dropped a **kidney**". (pick up and put on child).

Now let's put our hands on our **stomachs** (assistance if needed). Good, you know where it is.

Well, it looks like you learned a lot from this. Thanks so much for helping. **(make okay sign)**

Now we can **go back to class/watch a short video about digestion**.

<u>Digestion Script (Version 2 Game)</u>

Today, we're going to play a game and watch a video about digestion. First, let's **play a game/watch** a video.

There are lots of different types of food. Some are healthy and good for us, and some are not. We have a **hamburger and a piece of bread** (show items). These are just a few types of food we can eat when we're hungry.

The first thing we need to do is get food into our stomach right? Do you think we could swallow these **this piece of bread like this** (show bread) like this? Nooooo. That's why we need to chew up food before we swallow it...to help it fit in our stomach.

Look at these teeth (teeth model), these help break up our food into tiny pieces.

Let's help the big teeth brush. (Brush top teeth). There, I brushed the **top part** of the big teeth.

We also make saliva in our mouth – it's made right here (show salivary glands on wall chart). Saliva helps make the tiny pieces of food slimy...so that they can slide down your throat better. Let's break up the bread into tiny and crush it with the **hammer**. Now let's pretend to add some saliva (mash up bread with **hammer on the plate** and add water **[blue water]** in a bowl). We'll use this **blue** water for saliva, and add the bread.

Now we mix up the bread with this **plastic** spoon.

"Oops I spilled the **water**...silly me".

Once food is in our stomach, strong acid help to dissolve it and break it down even more. This way, we can pick up the good stuff our body needs later on. Sometimes when our stomachs are dissolving food, it can produce gas. That's why we burp. Let's put a bit more water and a fizzer into the **pickle jar** to make gas (put water and small bit of alka seltzer into jar). See how the gas is bubbling.

Our tongues help us to **taste the food**.

"I'm so hungry, I could eat a **house**"

Um, yummy cookie. (eat tiny bit of cookie)

Once we've got all the good stuff, we will have **energy for thinking.**

Did you know that **sheep** have more than one stomach to help them digest their food? Look at Danny the digester. This is where everything is happening (show on foam model) on Danny the Digester.

What's your favourite food Danny? (mouth with Danny) "My favourite food is **popcorn**".

It's important to wait for your food to digest **before you go to bed**.

When the acid in our stomachs has broken all the food down, it travels through our squiggly intestines down below (show on wall chart). Our squiggly intestines are big…mine would be as **long as a boat**. This is where our body picks up all the vitamins and good stuff we need to grow.

Some of the things we can't digest, we must filter out of our bodies. That's what the kidneys do – they filter bad stuff out of our bodies. These are your kidneys (show on chart). Let's add this **coffee** to this cup of water. See how it's all mixed together? Now when we pour it through this filter, all the **coffee** stays…and the water goes through (kids pour water and sand mix through filter). That's just like what our kidneys do – they filter out the bad stuff.

Down here is our tiny appendix. It's as **big as a finger**. It doesn't really do much digesting.

Now let's put on our digestion aprons. Here are all the organs we need to digest things. Let's try and place the organs in the right place our aprons. So let's put on the aprons here (assistance if needed putting aprons on).

First we'll start with the stomach…where does the stomach go (experimenter places **stomach** on apron). Oops, I put the **stomach on upside down**. (fix) Now let's put on the squiggly intestines. (Assistance). Now let's put the kidneys on (same).

Now let's put on the lungs so we can breathe.

These lungs look **wings from an angel**. (Drop lung) "Opps, **I dropped a lung**". (pick up and put on child).

Now let's put our hands on our **stomachs** (assistance if needed). Good, you know where it is.

Well, it looks like you learned a lot from this. Thanks so much for helping. **(make okay sign)**

Now we can **go back to class/watch a short video about digestion**.

<u>Digestion Script (Version 2 Video)</u>

Today, we're going to play a game and watch a video about digestion. First, let's **play a game/watch** a video.

There are lots of different types of food. Some are healthy and good for us, and some are not. Here's a banana **(use the banana as a guitar).** Look at me being silly using the banana as a **guitar.** Here are some other fruit **(show two grapes).** These are **grapes**. They're healthy for us. We also have **a box of juice** (show item). It's also healthy. These are just a few types of food we can eat when we're hungry.

The first thing we need to do is get food into our stomach right? Do you think we could swallow these **round crackers** (show crackers) like this? Nooooo. That's why we need to chew up food before we swallow it…to help it fit in our stomach.

Look at these teeth (teeth model), these help break up our food into tiny pieces. **"Goodbye"** (use teeth to say goodbye).Funny that the big teeth would be saying **goodbye**. See the teeth take a bite of food **(use teeth to take bite of banana).** Uh oh, the big teeth took a bite out of the **banana.**

Let's help the big teeth brush. (Brush **top** teeth). There, I brushed the **top part** of the big teeth.

Uh oh, it looks like Mr. teeth still has some **cavities** (show cavity).

We also make saliva in our mouth – it's made right here (show salivary glands on wall chart). Saliva helps make the tiny pieces of food slimy…so that they can slide down your throat better. Let's break up these **round** crackers with **the hammer** in this **clear bowl.**

Now we mix up the crackers with this **metal** spoon.

"Oops I spilled the **water**…silly me".

"I'm so hungry, I could eat a **house**"

What's your **favourite drink**?

Sometimes it hurts when our stomachs are **too full**?

Once food is in our stomach, strong acid help to dissolve it and break it down even more. This way, we can pick up the good stuff our body needs later on. Sometimes when our stomachs are dissolving food, it can produce gas. That's why we burp. Let's put a bit more water and a fizzer into this **pickle jar** to make gas (put water and small bit of alka seltzer into stomach). See how the gas is bubbling in the **pickle jar**. (make **squishy** noise). "**Squishy, squishy**….the food is digesting"

"Oh, the water is **hot**".

Once we've got all the good stuff, we will have **energy for playing.**

Did you know that **sheep** have more than one stomach to help them digest their food?

Look at Danny the digester. This is where everything is happening (show on foam model) on Danny the Digester.

I'll just take out Danny's **stomach**…there, I took out Danny the Digester's **stomach.**

It's important to wait for your food to digest **before you swim**.

Some of the things we can't digest, we must filter out of our bodies. That's what the kidneys do – they filter bad stuff out of our bodies. These are your kidneys (show on chart). Let's add this **dirt** to this cup of water. See how it's all mixed together? Now when we pour it through this filter, all the **dirt** stays…and the water goes through (kids pour water and sand mix through filter). That's just like what our kidneys do – they filter out the bad stuff.

Down here is our tiny appendix. It's as **big as a finger**. It doesn't really do much digesting.

Now let's put on our digestion aprons. Here are all the organs we need to digest things. Let's try and place the organs in the right place our aprons. So let's put on the aprons here (assistance if needed putting aprons on).

First we'll start with the stomach…where does the stomach go (experimenter places stomach on upside down on apron). Oops, I put the **stomach on upside down**. (Fixes it). Now let's put on the squiggly intestines. (Assistance). Now let's put the kidneys on (same).

Now let's put on the lungs so we can breathe. These lungs look **leaves from a tree**.

Now let's put our hands on our **squiggly intestines** (assistance if needed). Good, you know where it is.

Well, it looks like you learned a lot from this. Thanks so much for helping. **(make okay sign)**

Now we can **go back to class/watch a short video about digestion**.

<u>Version 1a (A)</u>

<u>Training:</u> Ok, you've done a good job at telling me about the things you learned about in the real-life game and about the things you learned about in the video.

<u>Control:</u> Ok, you've done a good job at telling me about the things you learned.

Now I'm going to ask you about some of the things you learned about last week with Sean.
I don't know exactly what happened that day, so I'm going to ask you some questions about it. I only want you to think about things that you learned that day. Okay? If you don't know the answer, it's OK to say "I don't know".

Note: After every fifth response, the interview will encourage the child by saying "good job". Where indicated, follow-up source questions ("Was that in the video or real-life game") will be asked if a child answers "yes" to the initial recognition question.

1. In the real life game, was there orange water mixed with the crackers?
2. In the video, did someone pretend the banana was a telephone?
3. In the video, was there a box of juice on the table?
4. In the real life game, was there a box of cereal on the table?
5. In the real life game, did someone use round crackers?
6. In the video, were the crackers mixed in the clear dish?
7. In the video, did the giant teeth say goodbye?
8. In the real life game, was there a plastic spoon on the table?
9. In the real life game, did someone tell you that your tongue helps you to swallow food?
10. In the video, did someone say that they were so hungry that they could eat a house?
11. Did someone show you how empty the stomach can be in the real life game?
12. In the video, did someone tell you that our bodies send more acid to our stomachs when we eat?
13. Did someone tell you that panda bears only eat one type of food in the real life game?
14. In the video, did you learn that cows have more than one stomach?
15. Did someone pour coffee through the filter in the video?
16. Remember when someone showed you how gas fizzes in the real life game, did they show you this in a pickle jar?

1. Were there two grapes on the table? (video or real life game)
2. Was there a hamburger on the table? (real life game or video)
3. Did someone use a hammer to crush the crackers? (video or real life game)
4. Did the big teeth take a bite of banana? (real life game or video)
5. Did someone take a bite of a cracker? (video or real life game)
6. Did someone ask you what your favourite drink was? (real life game or video)
7. Did someone make a squishy noise with the stomach? (real life game or video)
8. Did you learn that it hurts when you're full? (video or real life game)
9. Did Danny the Digester say that he likes popcorn? (real life game or video)
10. Did someone take out Danny the Digester's stomach? (video or real life game)
11. Did someone drop a lung on the floor? (real life game or video)

12. Did someone say that the lungs on the apron look like wings from an angel? (video or real life game)
13. Did someone put the stomach upside down on the apron? (real life game or video)
14. Did someone ask you to put your hands on your kidneys? (video or real life game)
15. Did someone make an okay sign at the end of the lesson? (real life game or video)
16. Was someone wearing a red shirt under his lab coat? (video or real life game)

"OK, we're all done now. You did great! Thanks so much for helping me. Let's go back to class now".

Version 1b (B)

Training: Ok, you've done a good job at telling me about the things you learned about in the real-life game and about the things you learned about in the video.

Control: Ok, you've done a good job at telling me about the things you learned.

Now I'm going to ask you about some of the things you learned about last week with Sean.
I don't know exactly what happened that day, so I'm going to ask you some questions about it. I only want you to think about things that you learned that day. Okay? If you don't know the answer, it's OK to say "I don't know".

Note: After every fifth response, the interview will encourage the child by saying "good job". Where indicated, follow-up source questions ("Was that in the video or real-life game") will be asked if a child answers "yes" to the initial recognition question.

17. Were there two grapes on the table? (video or real life game)
18. Was there a hamburger on the table? (real life game or video)
19. Did someone use a hammer to crush the crackers? (video or real life game)
20. Did the big teeth take a bite of banana? (real life game or video)
21. Did someone take a bite of a cracker? (video or real life game)
22. Did someone ask you what your favourite drink was? (real life game or video)
23. Did someone make a squishy noise with the stomach? (real life game or video)
24. Did you learn that it hurts when you're full? (video or real life game)
25. Did Danny the Digester say that he likes popcorn? (real life game or video)
26. Did someone take out Danny the Digester's stomach? (video or real life game)
27. Did someone drop a lung on the floor? (real life game or video)
28. Did someone say that the lungs on the apron look like wings from an angel? (video or real life game)
29. Did someone put the stomach upside down on the apron? (real life game or video)
30. Did someone ask you to put your hands on your kidneys? (video or real life game)
31. Did someone make an okay sign at the end of the lesson? (real life game or video)
32. Was someone wearing a red shirt under his lab coat? (video or real life game)

17. In the real life game, was there orange water mixed with the crackers?
18. In the video, did someone pretend the banana was a telephone?
19. In the video, was there a box of juice on the table?
20. In the real life game, was there a box of cereal on the table?
21. In the real life game, did someone use round crackers?
22. In the video, were the crackers mixed in the clear dish?
23. In the video, did the giant teeth say goodbye?
24. In the real life game, was there a plastic spoon on the table?
25. In the real life game, did someone tell you that your tongue helps you to swallow food?
26. In the video, did someone say that they were so hungry that they could eat a house?
27. Did someone show you how empty the stomach can be in the real life game?
28. In the video, did someone tell you that our bodies send more acid to our stomachs when we eat?
29. Did someone tell you that panda bears only eat one type of food in the real life game?
30. In the video, did you learn that cows have more than one stomach?
31. Did someone pour coffee through the filter in the video?
32. Remember when someone showed you how gas fizzes in the real life game, did they show you this in a pickle jar?

"OK, we're all done now. You did great! Thanks so much for helping me. Let's go back to class now".

Version 2a (C)

Training: Ok, you've done a good job at telling me about the things you learned about in the real-life game and about the things you learned about in the video.

Control: Ok, you've done a good job at telling me about the things you learned.

Now I'm going to ask you about some of the things you learned about last week with Sean.
I don't know exactly what happened that day, so I'm going to ask you some questions about it. I only want you to think about things that you learned that day. Okay? If you don't know the answer, it's OK to say "I don't know".

Note: After every fifth response, the interview will encourage the child by saying "good job". Where indicated, follow-up source questions ("Was that in the video or real-life game") will be asked if a child answers "yes" to the initial recognition question.

1. In the video, was there blue water mixed with crackers?
2. Was there a carton of milk on the table in the real life game?
3. Was there a piece of bread on the table in the video?
4. Did he use square crackers in the real life game?

5. In the real life game, were the crackers mixed in a clear dish?
6. In the real life game, did the giant teeth say goodbye?
7. Remember learning how gas fizzes in the real life game, did someone show you this in a mayonnaise jar?
8. In the real life game, did someone brush the bottom part of the giant teeth?
9. In the real life game, was there a metal spoon on the table?
10. In the video, did someone tell you that your tongue helps you swallow food?
11. In the video, did someone say that they were so hungry they could eat a car?
12. In the video, did you learn that cows have more than one stomach?
13. In the video, did you learn that it's good to wait for your food to digest before you go to bed?
14. In the video, did you learn that a squiggly intestine could be as long as a bus?
15. Did someone pour dirt through the filter in the real life game?
16. In the video, did someone tell you that we need food energy for thinking?

1. Was there a hamburger on the table? (real life game or video)
2. Did you learn that the big teeth have cavities? (video or real life game)
3. Did someone spill the crackers? (real life game or video)
4. Did someone take a bite of the bread? (video or real life game)
5. Did someone ask you what your favourite food was? (real life game or video)
6. Did someone say that the water in the jar was hot? (video or real life game)
7. Did someone make a squishy noise with the stomach? (video or real life game)
8. Did Danny the Digester say that he likes popcorn? (video or real life game)
9. Did someone take out Danny the Digester's squiggly intestine? (real life game or video)
10. Remember the appendix was the little thing at the bottom of the stomach. Did you learn that it was as big as a pencil? (video or real life game)
11. Did someone drop a kidney on the floor? (real life game or video)
12. Did someone say that the lungs look like wings from an angel? (video or real life game)
13. Did someone put the squiggly intestines upside down on the apron? (real life game or video)
14. Did someone ask you to put your hands on your kidneys? (video or real life game)
15. Did someone give a high five at the end of the lesson? (real life game or video)
16. Was someone wearing a grey shirt under his lab coat? (video or real life game)

"OK, we're all done now. You did great! Thanks so much for helping me. Let's go back to class now".

<u>Version 2b (D)</u>

<u>Training:</u> Ok, you've done a good job at telling me about the things you learned about in the real-life game and about the things you learned about in the video.

<u>Control:</u> Ok, you've done a good job at telling me about the things you learned.

Now I'm going to ask you about some of the things you learned about last week with Sean.
I don't know exactly what happened that day, so I'm going to ask you some questions about it. I only want you to think about things that you learned that day. Okay? If you don't know the answer, it's OK to say "I don't know".

Note: After every fifth response, the interview will encourage the child by saying "good job". Where indicated, follow-up source questions ("Was that in the video or real-life game") will be asked if a child answers "yes" to the initial recognition question.

17. Was there a hamburger on the table? (real life game or video)
18. Did the big teeth have cavities? (video or real life game)
19. Did someone spill the crackers? (real life game or video)
20. Did someone take a bite of the bread? (video or real life game)
21. Did someone ask you what your favourite food was? (real life game or video)
22. Did someone say that the water in the jar was hot? (video or real life game)
23. Did someone make a squishy noise with the stomach? (video or real life game)
24. Did Danny the Digester say that he likes popcorn? (video or real life game)
25. Did someone take out Danny the Digester's squiggly intestine? (real life game or video)
26. Remember the appendix was the little thing at the bottom of the stomach. Did you learn that it was as big as a pencil? (video or real life game)
27. Did someone drop a kidney on the floor? (real life game or video)
28. Did someone say that the lungs look like wings from an angel? (video or real life game)
29. Did someone put the squiggly intestines upside down on the apron? (real life game or video)
30. Did someone ask you to put your hands on your kidneys? (video or real life game)
31. Did someone give a high five at the end of the lesson? (real life game or video)
32. Was someone wearing a grey shirt under their lab coat? (video or real life game)
17. In the video, was there blue water mixed with the crackers?
18. Was there a carton of milk on the table in the real life game?
19. Was there a piece of bread on the table in the video?
20. Did someone use square crackers in the real life game?
21. In the real life game, were the crackers mixed in the clear dish?
22. In the real life game, did the giant teeth say goodbye?
23. In the real life game, did someone brush the bottom part of the giant teeth?
24. In the real life game, was there a metal spoon on the table?
25. Remember learning how gas fizzes in the real life game, did someone show you this in a mayonnaise jar?
26. In the video, did someone tell you that your tongue helps you to swallow food?
27. In the video, did someone tell you that we need food energy for thinking?
28. In the video, did someone say that they were so hungry they could eat a car?
29. In the video, did you learn that cows have more than one stomach?

30. In the video, did you learn that it's good to wait for your food to digest before you go to bed?
31. In the video, did you learn that a squiggly intestine could be as long as a bus?
32. Did someone pour dirt through the filter in the real life game?

"OK, we're all done now. You did great! Thanks so much for helping me. Let's go back to class now".

Appendix B: Frog Training Script and Questions (Study 1)

FROG SCRIPT - Version 1 Game (Version 2 in Video)

Invitation:
"Hi, my name is [name] I heard that someone called [name] came to school last week. And I heard that you learned about digestion with her. "Today we are going to learn about frogs, there are lots of neat things we can learn about frogs Let's go sit on the mat."

First we are going to watch a **video/ play a game**.

Event:

Video only: "Hello, and welcome to all about frogs."
"Now we are going to put on our frog eyes so we can learn about frogs" [give each child a pair of frog eyes to put on]

A really interesting thing about frogs is how they grow up. Frogs are born in the water like fish. They are called tadpoles. This is Timmy the tadpole [show tadpole]. Let's watch Timmy the tadpole swim [put tadpole in water and make swim].
Timmy likes to swim in the **blue water**. "Splash, splash, splash!" Look as Timmy swims by the **white flower** on the lily pad.

When Timmy grows up he will grow legs and turn into a frog. This is where his legs grow from [show legs on life cycle poster].

When frogs are all grown up they can also live out of the water like Frankie the Frog. [show Frankie] Say Hello to Frankie.

Let's put a sticker of a frog like Frankie the frog in the pond. [Have children place sticker on pond picture]

Frankie likes to **play hide and seek**. Look at the other frog chase Frankie as they play hide and seek. [have another frog chase Frankie]

Look at Frankie's sticky tongue. Do you want to feel how sticky his tongue is? [have kids touch tongue]

Frankie uses his tongue to catch bugs. Let's watch Frankie **catch the fly** with his tongue.

Frankie's tongue can stretch as big as a **pencil**. [place pencil next to tongue].

Frankie has to be careful that other animals don't eat him! **Racoons** like to chase Frogs. Racoon puppet –He,he..look at me chase the frog. [Puppet chases frog]. Uh-Oh, The Racoon chased Frankie the Frog.

One of the biggest types of frogs is a **Bullfrog**. Bullfrogs are really big frogs just like Bulls are really big. This is what a Bullfrog sounds like [push button on Chicco].

The biggest frog in the world is as big as a toy **truck**. [Show toy truck]. That is pretty big isn't it!!!

There are many different types of frogs. These are called poison dart frogs and they have lots of different colours on them. [show small plastic poison dart frog to kids]

Before we finish, let's put the frogs in the pond. [have children place fantastic frogs into the aquarium]

<u>FROG SCRIPT - Version 2 Game (Version 1 in video)</u>

<u>Invitation:</u>
"Hi, my name is [name] I heard that someone called [name] came to school last week. And I heard that you learned about digestion with her. "Today we are going to learn about frogs, there are lots of neat things we can learn about frogs Let's go sit on the mat."

First we are going to **play a game/watch a video.**

<u>Event:</u>

Video only: "Hello, and welcome to all about frogs."
"Now we are going to put on our frog eyes so we can learn about frogs" [give each child a pair of frog eyes to put on]

A really interesting thing about frogs is how they grow up. Frogs are born in the water like fish. They are called tadpoles. This is Timmy the tadpole [show tadpole]. Let's watch Timmy the tadpole swim [put tadpole in water and make swim].
Timmy likes to swim in the **green water**. "Splash, splash, splash!" Look as Timmy swims by the **yellow flower** on the lily pad.

When Timmy grows up he will grow legs and turn into a frog. This is where his legs grow from [show legs on life cycle poster].

When frogs are all grown up they can also live out of the water like Frankie the Frog. [show Frankie] Say Hello to Frankie.

Let's put a sticker of a frog like Frankie the frog in the pond. [Have children place sticker on pond mat]

Frankie likes to play **tag**. Look at the other frog chase Frankie as they play tag. [have another frog chase Frankie]

Look at Frankie sitting on the log. Frankie likes to sit on the **log**

Look at Frankie's sticky tongue. Do you want to feel how sticky his tongue is? [have kids touch tongue]

Frankie's tongue can stretch as big as a **spoon**. [place pencil/spoon next to tongue].

Frankie has to be careful that other animals don't eat him! **Fish** like to chase Frogs. Fish puppet –He,he..look at me chase the frog. [Puppet chases frog]. Uh-Oh, The Fish chased Frankie the Frog.

Frogs like to sleep in the **mud in the winter**. This is called hibernation. [Show mud on chart or in aquarium frog in mud] This makes them all muddy! Yuck!

The biggest frog in the world is as big as a **telephone**. [Show telephone]. That is pretty big isn't it!!!

There are many different types of frogs. These are called poison dart frogs and they have lots of different colours on them. [show small plastic poison dart frog to kids]

Before we finish, let's put the frogs in the pond. [have children place fantastic frogs around pond mat]

Control Training Questions

1. Did the frog catch a fly with his tongue?

Positive reinforcement

2. Did the frog stay in the mud during the winter? (novel)

Positive reinforcement

3. Did a Racoon chase the frog?

Positive reinforcement

4. Think about the biggest frog. Was the frog as big as a "Toy Truck"?

Positive reinforcement

Set Training Questions –Version 1

1. Did the frog catch a fly with his tongue?

A) Did the frog catch a fly with his tongue in the real life game or in the video?

Positive reinforcement

2. Did the frog stay in the mud during the winter?
   A) Did the frog stay in the mud in the video or in the real life game?

Positive reinforcement

3. Did a Racoon chase the frog?
   A) Did the Racoon chase the frog in the real life game or the video?

Positive reinforcement

4. Think about the biggest frog. Was the frog as big as a Telephone"?
   A) Was the telephone in the video or real life game?

Positive reinforcement

It looks like you're good at thinking about where you learned things from.

<u>Set Training Questions –Version 2</u>

1. Did the frog catch a fly with his tongue?
   B) Did the frog catch a fly with his tongue in the real life game or in the video?

Positive reinforcement

2. Did the frog stay in the mud during the winter?
   B) Did the frog stay in the mud in the video or in the real life game?

Positive reinforcement

3. Did a Racoon chase the frog?
   B) Did the Racoon chase the frog in the real life game or the video?

Positive reinforcement

4. Think about the biggest frog. Was the frog as big as a Telephone"?
   B) Was the telephone in the video or real life game?

Positive reinforcement

It looks like you're good at thinking about where you learned things from.

<u>Criterion Training Questions –Version 1</u>

*No correction of recognition accuracy (only correction of source accuracy) on initial 10 questions. If incorrect, move on to next question.

*If child has not reached criterion after 10 questions, repeat questions with recognition correction AND source correction.

1. Did the frog catch a fly with his tongue?
   C) Did the frog catch a fly with his tongue in the real life game or in the video?
(yes in real life game)

If video, "That's a good guess, but actually, the frog caught the fly with his tongue in the real life game"

Positive reinforcement

2. Did the frog stay in the mud during the winter?
   C) Did the frog stay in the mud in the video or in the real life game?
(yes in video)

If real life game, "That's a good guess, but actually, the frog stayed in the mud in the video"

Positive reinforcement

3. Did a Racoon chase the frog?
   C) Did the Racoon chase the frog in the real life game or the video?

(yes in real life game)

If video, "That's a good guess, but actually, the Racoon chased the frog in the real life game"

Positive reinforcement

4. Think about the biggest frog. Was the frog as big as a Telephone"?
   C) Was the telephone in the video or real life game?
(yes in video)

If real life game, "That's a good guess, but actually, the telephone was in the video"

Positive reinforcement

5. Did "Frankie the frog" like playing hide and seek?

A) Did Frankie like playing hide and seek in the real life game or the video?
(yes in real life game)

If video, "That's a good guess, but actually, he played hide and seek in the real life game"

Positive reinforcement

6. Did you learn about a big bullfrog?
A) Did you learn about big bullfrogs in the video or the real life game?

(yes in real life game)

If video, "That's a good guess, but actually, you learned about bullfrogs in the real life game"

Positive reinforcement

7. Did you learn that a frog's tongue is as long as a pencil?
A) Was the pencil in the real life game or the video?
(yes in real life game)

If video, "That's a good guess, but actually, the pencil was in the real life game"

8. Was a tadpole swimming in green water?
A) Was the tadpole swimming in green water in the video or the real life game?

(yes in video)

If real life game, "That's a good guess, but actually, the tadpole was swimming in the green water in the video".

Positive reinforcement

9. Was there a white flower on a lily pad?

A) Was the white flower in the real life game or the video?

(yes in real life game)

If video, "That's a good guess, but actually, the white flower was in the real life game"

Positive reinforcement

10. Was the frog sitting on a log?

A) Was the frog sitting on a log in the video or the real life game?

(yes in video)

If real life game, "That's a good guess, but actually, the frog was sitting on a log in the video"

Positive reinforcement

At the end of Criterion

It looks like you're good at thinking about where you learned things from.

Criterion Training Questions –Version 2

*No correction of recognition accuracy (only correction of source accuracy) on initial 10 questions. If incorrect, move on to next question.

*If child has not reached criterion after 10 questions, repeat questions with recognition correction AND source correction.

1. Did the frog catch a fly with his tongue?
   D) Did the frog catch a fly with his tongue in the real life game or in the video?

(yes in video)

If real life game, "That's a good guess, but actually, the frog caught the fly with his tongue in the video"

Positive reinforcement

2. Did the frog stay in the mud during the winter?
   D) Did the frog stay in the mud in the video or in the real life game?

(yes in real life game)

If video, "That's a good guess, but actually, the frog stayed in the mud in the real life game"

Positive reinforcement

3. Did a Racoon chase the frog?
   D) Did the Racoon chase the frog in the real life game or the video?

(yes in video)

If real life game, "That's a good guess, but actually, the Racoon chased the frog in the video"

Positive reinforcement

4. Think about the biggest frog. Was the frog as big as a Telephone"?
    D) Was the telephone in the video or real life game?
(yes in real life game)

If video, "That's a good guess, but actually, the telephone was in the real life game"

Positive reinforcement

5. Did "Frankie the frog" like playing hide and seek?
    B) Did Frankie like playing hide and seek in the real life game or the video?
(yes in video)

If real life game, "That's a good guess, but actually, he played hide and seek in the video"

Positive reinforcement

7. Did you learn about a big bullfrog? (novel)
    B) Did you learn about big bullfrogs in the video or the real life game?

    (yes in video)

    If real life game, "That's a good guess, but actually, he played hide and seek in the video"

Positive reinforcement

7. Did you learn that a frog's tongue is as long as a pencil?
    B) Was the pencil in the real life game or the video?
(yes in video)

If real life game, "That's a good guess, but actually, the pencil was in the video"

8. Was a tadpole swimming in green water?
    B) Was the tadpole swimming in green water in the video or the real life game?

    (yes in real life game)

If video, “That’s a good guess, but actually, the tadpole was swimming in the green water in the real life game”.

Positive reinforcement

9. Was there a white flower on a lily pad?

B) Was the white flower in the real life game or the video?

(yes in video)

If real life game, “That’s a good guess, but actually, the white flower was in the video”

Positive reinforcement

10. Was the frog sitting on a log?

B) Was the frog sitting on a log in the video or the real life game?

(yes in real life game)

If video, “That’s a good guess, but actually, the frog was sitting on a log in the real life game”

Positive reinforcement

At the end of Criterion

It looks like you’re good at thinking about where you learned things from.

Appendix C: Nature Training Script and Questions (Study 2)

Version 1 Nature Script (Game)

OK – so I as mentioned to you before, we're going to be learning all about nature. First we're going to watch a video about nature, and then we're going do a real life activity about nature (presentation order balanced).

There are 3 really important things that make up what we call nature (use fingers to show): 1) the weather 2) plants and 3) animals. All of these things play very important roles in our earth – and all have very cool features about them. Let me move **this box** out of the way so you can see.

First let's talk about weather. Often when people think of weather, they think of wind. Sometimes when it's really windy, a tornado can happen. A tornado is when the wind spins really fast in a circle. **I'll show you how this works with this pop bottle. Watch the swirly water in the pop bottle.** The wind is important to nature. It helps to **carry seeds** so that they can be planted away from the trees they came from. Sometimes when it is really stormy out, it rains. This rain helps to water the plants – because it trickles down through the ground to the roots of the plants. Watch as I pour the water **through the rocks** (pour water). If there were roots down there, they could pick that water up. **Some plants don't like very much water – like cactuses (show picture). Cactuses don't need very much water.** Sunlight helps to warm the planet and dry up some of the water. My favourite **season is summer, because I like the sun.**

Now let's talk about plants. Plants are very important to us. They give us oxygen, they give us food, and they help to clean the air that we breathe. Some plants are good to eat, and others are not. We even get some medicine from plants – **like Aspirin, Aspirin comes from plants** – did you know that? Plants use their roots to pick up water and nutrients from the soil to help them live. The water flows up the root against gravity because it sticks to itself really well. Watch this **blue water** travel up the paper towel (show). Even though I haven't put this part of the towel in water, it's now wet. That's because the water travelled up by itself. Some roots we eat. Did you know that the **radishes we eat are roots? This is a radish (show radish).** This is a seed. Let's plant it in the pot of soil. Oops, **I spilled some of the soil.** Now if we want this plant to grow, we'd have to keep the soil wet. This is a **red leaf**. Plants use leaves to absorb the sunlight they need to live. Some plants have lots of air in them. This is an **apple, it has very little air in it.**

Now let's talk a bit about animals. They are also very important to our earth. Different animals find food in different ways. **This is a Racoon. He's a scavenger – which means that he finds food lying around.** His name is Ricky – **say hello Ricky (puppet says hello). Do you want to feel how soft the raccoon is?** Other animals hunt for their food. To hunt for food, you have to be very fast. Did you know that **a cheetah is the fastest land animal** in the world (show picture)? They are very good at catching their prey. Some animals use poison to catch their prey, and others use their tongues. This is a frog (show frog). His tongue can stretch as **long as a pencil.** He uses this to catch bugs. **Did you know that frogs change throughout their lifespan? They (show pictures) change from a tadpole to a frog.** Do know what this is (show picture)? This is **a whale**– they live under the water. **What's your favourite animal?**

Now ~~let's~~ I'm going to quickly draw a picture. ~~We~~ I don't have much time, so ~~we'll~~ I'll have to make it a very quick drawing. Here's some **crayons** ~~we~~ I can use to draw with. I'm going to draw **a cloudy day.** (begin drawing). Look, this cloud looks like **a flower** doesn't it? This is a **cathouse so my cat doesn't get wet.** I'll also draw some **very short grass – it looks like someone just cut this grass.**

Now, can you help me put a few of these things away (pick a few items)? Thanks! That's very helpful.

<u>Version 1 Nature Script (Video)</u>

OK – so I as mentioned to you before, we're going to be learning all about nature. First we're going to watch a video about nature, and then we're going do a real life activity about nature (presentation order balanced).

There are 3 really important things that make up what we call nature (use fingers to show): 1) the weather 2) plants and 3) animals. All of these things play very important roles in our earth – and all have very cool features about them. Let me move **this table** out of the way so you can see.

First let's talk about weather. Often when people think of weather, they think of wind. Sometimes when it's really windy, a tornado can happen. A tornado is when the wind spins really fast in a circle. The wind is important to nature. It helps **leaves fall from the trees**. Sometimes when it is really stormy out, it rains. This rain helps to water the plants – because it trickles down through the ground to the roots of the plants. Watch as I pour the water **through the dirt** (pour water). **This is a rain gauge, it is used to measure how much rain we get.** If there were roots down there, they could pick that water up. **Some plants like a lot of water – like moss (show picture). Moss needs dark, damp environments.** Sunlight helps to warm the planet and dry up some of the water. My favourite **season is winter, because I like the snow.**

Now let's talk about plants. Plants are very important to us. They give us oxygen, they give us food, and they help to clean the air that we breathe. Some plants are good to eat, and others are not. We even get some medicine from plants – **like Antibiotics, Antibiotics come from plants –** did you know that? Plants use their roots (show picture) to pick up water and nutrients from the soil to help them live. The water flows up the root against gravity because it sticks to itself really well. Watch this **green water** travel up the paper towel (show). Even though I haven't put this part of the towel in water, it's now wet. That's because the water travelled up by itself. Some roots we eat. Did you know that the **carrots we eat are roots? This is a carrot (show picture).** Here's a seed. Let's plant it in the pot of soil. Now if we want this plant to grow, we'd have to keep the soil wet. We can use this **moisture meter** to test if the soil is wet enough. **This is a germinated grass seed – this is what happens when it begins to grow**. This is a **yellow leaf.** Plants use leaves to absorb the sunlight they need to live. Some plants have lots of air in them. This is **a potato, it has very little air in it – which makes it sink in water. What's your favourite plant?**

Now let's talk a bit about animals. They are also very important to our earth. Different animals find food in different ways. **This is a Fish. Fish are hunters – which**

means that they chase their food. His name is **Freddy – say goodbye Freddy (puppet says goodbye).** To hunt for food, you have to be very fast. Did you know that **a Peregrine Falcon is the fastest bird** in the world (show picture)? They are very good at catching their prey. Some animals use poison to catch their prey, and others use their tongues. This is a frog (show frog). His tongue can stretch as **long as a spoon.** He uses this to catch bugs. Do know what this is (show picture)? This is **a turtle** – they live under the water. Some animals hibernate during the winter – **like mice, mice hibernate until it gets warm again.**

Now let's quickly draw a picture. We don't have much time, so we'll have to make it a very quick drawing. Here's some **paint** we can use to draw with. I'm going to draw **a sunny day.** (begin drawing). Look, this cloud looks like **a heart** doesn't it? This is a **doghouse so my dog has somewhere to go.** I'll also draw some **very tall grass – it looks like someone needs to cut this grass. Do you think my picture has enough grass?**

Version 2 Nature Script (Game)

OK – so I as mentioned to you before, we're going to be learning all about nature. First we're going to watch a video about nature, and then we're going do a real life activity about nature (presentation order balanced).

There are 3 really important things that make up what we call nature (use fingers to show): 1) the weather 2) plants and 3) animals. All of these things play very important roles in our earth – and all have very cool features about them. Let me move **this chair** out of the way so you can see. **Can you help me take a few of the items out?** Thanks, you're very helpful.

First let's talk about weather. Often when people think of weather, they think of wind. Sometimes when it's really windy, a tornado can happen. A tornado is when the wind spins really fast in a circle. **I'll show you how this works with this water bottle. Watch the swirly water in the water bottle.** The wind is important to nature. It helps **to create waves in the ocean which help underwater plants and animals. This is a wind gauge; we use it to measure the speed of wind.** Have you ever been outside **on a windy day?** Sometimes when it is really stormy out, it rains. This rain helps to water the plants – because it trickles down through the ground to the roots of the plants. ~~**This is a rain gauge, it is used to measure how much rain we get.**~~ Sunlight helps to warm the planet and dry up some of the water. Also when it's stormy out, you can sometimes see lightening. Lightening comes from static electricity. See how the **pieces of paper stick** to the balloon (show), this is because of static electricity. Do you know what you're supposed to do when an **earthquake comes? You're supposed to hide under a table or something secure**.. What do you like to do in the **Winter when it's cold?**

Now let's talk about plants. Plants are very important to us. They give us oxygen, they give us food, and they help to clean the air that we breathe. Some plants are good to eat, and others are not. This is a picture of berries. These are **deadly berries** (show picture) – so we can't eat them. You should never eat anything you find in the wild. Plants use their roots (show picture) to pick up water and nutrients from the soil to help them live. The water flows up the root against gravity because it sticks to itself really well. Here's a seed. Let's plant it in the pot of soil. Oops, **I spilled some of the seeds.** First, we should cultivate the soil **with the rake**. Now if we want this plant to grow, we'd have to keep the soil at the right ph. The ph is the amount of acid in the soil. We can use this **ph meter** to test if the soil is acidic enough. **This is a germinated maple leaf seed – this is what happens when it begins to grow**. This is a **green leaf**. Plants use leaves to absorb the sunlight they need to live. Some plants have lots of air in them. This is **celery, it has lots of air in it – which makes it float in water.** Some plants eat bugs**. This is a Venus Fly Trap** (show picture)**. It eats bugs.**

Now let's talk a bit about animals. They are also very important to our earth. Different animals find food in different ways. **This is a Fox. Foxes are scavengers – which mean that they find food lying around.** His name is **Freddy. Do you want to feel how soft the fox is?** Some animals use poison to catch their prey, and others use their tongues. This is a frog (show frog). He uses his tongue to catch bugs. Other animals defend themselves with poison. **This is a poisonous snake** (show). Some animals hibernate during the winter – **like bears. Bears hibernate until it gets warm again**. Some Leopards (show model) like to **live in warm, wet jungles. Did you know that caterpillars change throughout their lifespan? They (show pictures) change from a caterpillar to a butterfly.** Let's look at **a spider** under this magnifying glass. Kind of cool isn't it?

Now ~~let's~~ I'm going to quickly draw a picture. We don't have much time, so we'll have to make it a very quick drawing. Here's some **markers** ~~**we can use**~~ I'm going to draw with. I'm going to draw **an animal.** (begin drawing). **Do you think my picture has enough trees?**

OK – I think our time is up. (Move on to next activity) ~~**We should get back to class now. Thanks for helping me learn all about nature. Let's go back to class now.**~~

Version 2 Nature Script (Video)

OK – so I as mentioned to you before, we're going to be learning all about nature. First we're going to watch a video about nature, and then we're going do a real life activity about nature (presentation order balanced).

There are 3 really important things that make up what we call nature (use fingers to show): 1) the weather 2) plants and 3) animals. All of these things play very important roles in our earth – and all have very cool features about them.

First let's talk about weather. Often when people think of weather, they think of wind. Sometimes when it's really windy, a tornado can happen. A tornado is when the wind spins really fast in a circle. **I'll show you how this works with this juice bottle.** Do you know what you're supposed to do when **a tornado comes? You're supposed to hide under a table or something secure**. Have you ever been outside **on a cloudy day?** Also when it's stormy out, you can sometimes see lightening. Lightening comes from static electricity. See how **my hair sticks** to the balloon (show), this is because of static electricity. Sometimes when it is really stormy out, it rains. This rain helps to water the plants – because it trickles down through the ground to the roots of the plants. Snow is also important to nature. **This is how we measure snow** (show ruler)**.** Sunlight helps to warm the planet and dry up some of the water. What do you like to do in the **Summer when it's warm?** My favourite **season is fall.**

Now let's talk about plants. Plants are very important to us. They give us oxygen, they give us food, and they help to clean the air that we breathe. Some plants are good to eat, and others are not. These are **deadly mushrooms** (show picture) – so we can't eat them. You should never eat anything you find in the wild. Plants use their roots (show picture) to pick up water and nutrients from the soil to help them live. The water flows up the root against gravity because it sticks to itself really well. Watch this **red water** travel up the paper towel (show). That's because the water travelled up by itself. Here's a seed. Let's plant it in the pot of soil. Oops, **I spilled the water.** First, we should cultivate the soil **with the shovel**. We can run our hands through the soil to test and see **if it's too rocky. This is a germinated flower seed – this is what happens when it begins to grow**. Plants use leaves to absorb the sunlight they need to live. Some plants have lots of air in them. Some plants eat bugs. **This is a Sundew Plant** (show picture)**. It eats bugs.** Kind of cool isn't it?

Now let's talk a bit about animals. They are also very important to our earth. Different animals find food in different ways. To hunt for food, you have to be very fast. Did you know that **a Sailfish is the fastest fish** in the world (show picture)? They are very good at catching their prey. Some animals use poison to catch their prey, and others use their tongues. This is a frog (show frog). He uses his tongue to catch bugs. Some animals hibernate during the winter – **like bats (show picture). Bats hibernate until it gets warm again**. Some animal defend themselves with poison. **This is a poisonous frog** (show). Some Leopards (show picture) like to **live in high cold mountains. Did you know that flies change throughout their lifespan? They (show pictures) change from a maggot to a fly.** Let's look at **a fly** under this magnifying glass (show).

Now let's quickly draw a picture. We don't have much time, so we'll have to make it a very quick drawing. Here's some **pens** we can use to draw with. I'm going to draw **a snowy day.** (begin drawing). **Do you think my picture has enough clouds?** Look, I drew **a birdhouse in my picture.**

OK – I think our time is up. We should get back to class now. Thanks for helping me learn all about nature. Can you **push that chair in before we go?** Thanks! Let's go back to class now.

Baseline question: Did you feel how soft the raccoon was, in the game or the video (counterbalance order)?
Baseline correct?: ____

Tier 1 Questions (NM T1)

Question	**Corr Ans**	**Source Ans**	**1st Src given – V or G**	**Child Respn.**	**R. Ac.**	**S. Ac.**	**Q Num**	**Feedbk Given?**	**Total # Corr. (tier)**	**Total # Incor (tier)**
Did you see swirly water in a pop bottle?	Yes	game								
Did someone show you how to measure rain?	Yes	video								
Did someone use crayons?	Yes	game								
Did you learn that mice like to hibernate?	Yes	video								
Did you learn that frogs change throughout their lifespan?	Yes	game								
Was there a germinated grass seed?	Yes	video								
Did someone spill the soil?	Yes	game								
Did you see how we test if soil is wet?	Yes	video								
Were you asked to help them put the items away?	Yes	game								
Did someone ask if their picture has enough grass?	Yes	video								

Move up – **2 consecutive correct**
Stop – **6 total correct at any one tier**
Move down – **2 consecutive incorrect**
Statement at 5 incorrect: **"Try to think hard about *where* you saw things"**
Statement at 7 incorrect: **"Remember you saw a video and a real life game, try to think of the differences between them"**
Maximum of 25 questions

Tier 2 Questions (M T1)

Question	**Corr. Answ.**	**Child Respon.**	**R. Ac.**	**S. Ac.**	**Q Num**	**Feedback Given?**	**Total # Corr. (tier)**	**Total # Incor (tier)**
Did you learn how static electricity makes a piece of paper stick to balloon?	No							
Did someone teach you what to do when a tornado comes?	No							
Was there a poisonous snake?	No							
Did you learn that some leopards live in high, cold mountains?	No							
Was there a spider under a magnifying glass?	No							
Was a shovel used to cultivate the soil?	No							
Did you learn that Venus flytraps eat bugs?	No							
Was there a picture of a deadly mushroom?	No							
Did someone ask you what you like to do in the summer?	No							
Were you asked if you've ever been outside on a windy day?	No							

Move up – **2 consecutive correct**
Stop – **6 total correct at any one tier**
Move down – **2 consecutive incorrect**

Tier 3 (NM T2)

Question	**Corr. Answ.**	**Src Answ.**	**1st Src given – V or G**	**Child Respn.**	**R. Ac.**	**S. Ac.**	**Q Num**	**Feedbk Given?**	**Total # Corr. (tier)**	**Total # Incor (tier)**
Was there water that trickled through rocks?	Yes	game								
Did you learn that moss likes a lot of rain?	Yes	video								
Did you learn that a frog's tongue is as big as a pencil?	Yes	game								
Did You learn that turtles live under the water?	Yes	video								
Did one of the animals say hello?	Yes	game								
Did you learn that carrots are a root we eat?	Yes	video								
Did you learn that aspirin comes from plants?	Yes	game								
Did someone draw very tall grass on their picture?	Yes	video								
Were you asked what your favourite animal is?	Yes	game								
Was there a cloud that looked like a heart?	Yes	video								

Move up – **2 consecutive correct**
Stop – **6 total correct at any one tier**
Move down – **2 consecutive incorrect**

Tier 4 (M T2)

Question	**Corr. Answ.**	**Child Respon.**	**R. Ac.**	**S. Ac.**	**Q Num**	**Feedback Given?**	**Total # Corr. (tier)**	**Total # Incor (tier)**
Did you learn that wind is important to create waves in the ocean?	No							
Did someone draw a snowy day?	No							
Did you learn about a fox as a scavenger?	No							
Did you learn that the sailfish is the fastest fish in the world?	No							
Was there a green leaf?	No							
Did you see red water go through the paper towel?	No							
Did you learn that celery has lots of air in it?	No							
Did someone say their favourite season is the fall?	No							
Was a chair moved out of the way so you could see better?	No							
Was there a birdhouse in one of the pictures?	No							

Move up – **2 consecutive correct**
Stop – **6 total correct at any one tier**
Move down – **2 consecutive incorrect**

Baseline question: Did you feel how soft the fox was, in the game or the video (counterbalance order)?
Baseline correct?: ____
Tier 1 (NM T1)

Question	Corr. Answ.	Src Answ.	1st Src given – V or G	Child Respn.	R. Ac.	S. Ac.	Q Num	Feedbk Given?	Total # Corr. (tier)	Total # Incor (tier)
Did you learn that wind is important to create waves in the ocean?	Yes	game								
Did someone draw a snowy day?	Yes	video								
Did you learn about a fox as a scavenger?	Yes	game								
Did you learn that the sailfish is the fastest fish in the world?	Yes	video								
Was there a green leaf?	Yes	game								
Did you see red water go through the paper towel?	Yes	video								
Did you learn that celery has lots of air in it?	Yes	game								
Did someone say their favourite season is the fall?	Yes	video								
Was a chair moved out of the way so you could see better?	Yes	game								
Was there a birdhouse in one of the pictures?	Yes	video								

Move up – **2 consecutive correct**
Stop – **6 total correct at any one tier**
Move down – **2 consecutive incorrect**
Statement at 5 incorrect: **"Try to think hard about *where* you saw things"**
Statement at 7 incorrect: **"Remember you saw a video and a real life game, try to think of the differences between them"**
Maximum of 25 questions

Tier 2 (M T1)

Question	**Corr. Answ.**	**Child Respon.**	**R. Ac.**	**S. Ac.**	**Q Num**	**Feedback Given?**	**Total # Corr. (tier)**	**Total # Incor (tier)**
Was there water that trickled through rocks?	No							
Did you learn that moss likes a lot of rain?	No							
Did you learn that a frog's tongue is as big as a pencil?	No							
Did You learn that turtles live under the water?	No							
Did one of the animals say hello?	No							
Did you learn that carrots are a root we eat?	No							
Did you learn that aspirin comes from plants?	No							
Did someone draw very tall grass on their picture?	No							
Were you asked what your favourite animal is?	No							
Was there a cloud that looked like a heart?	No							

Move up – **2 consecutive correct**
Stop – **6 total correct at any one tier**
Move down – **2 consecutive incorrect**

Tier 3 Questions (NM T2)

Question	**Corr. Answ.**	**Src. Answ.**	**1st src given – V or G**	**Child Respn.**	**R. Ac.**	**S. Ac.**	**Q Num**	**Feedbk Given?**	**Total # Corr. (tier)**	**Total # Incor (tier)**
Did you learn how static electricity makes a piece of paper stick to balloon?	Yes	Game								
Did someone teach you what to do when a tornado comes?	Yes	video								
Was there a poisonous snake?	Yes	game								
Did you learn that some leopards live in high, cold mountains?	Yes	video								
Was there a spider under a magnifying glass?	Yes	game								
Was a shovel used to cultivate the soil?	Yes	video								
Did you learn that Venus flytraps eat bugs?	Yes	game								
Was there a picture of a deadly mushroom?	Yes	video								
Did someone ask you what you like to do in the summer?	Yes	video								
Were you asked if you've ever been outside on a windy day?	Yes	game								

Move up – **2 consecutive correct**
Stop – **6 total correct at any one tier**
Move down – **2 consecutive incorrect**

Tier 4 Questions (M T2)

Question	**Corr. Answ.**	**Child Respon.**	**R. Ac.**	**S. Ac.**	**Q Num**	**Feedback Given?**	**Total # Corr. (tier)**	**Total # Incor (tier)**
Did you see swirly water in a pop bottle?	No							
Did someone show you how to measure rain?	No							
Did someone use crayons?	No							
Did you learn that mice like to hibernate?	No							
Did you learn that frogs change throughout their lifespan?	No							
Was there a germinated grass seed?	No							
Did someone spill the soil?	No							
Did you learn how we test if soil is wet?	No							
Were you asked to help them put the items away?	No							
Did someone ask if their picture has enough grass?	No							

Move up – **2 consecutive correct**
Stop – **6 total correct at any one tier**
Move down – **2 consecutive incorrect**

References

Ackil, J.K., & Zaragoza, M.S. (1995). Developmental differences in eyewitness suggestibility and memory for source. *Journal of Experimental Child Psychology, 60,* 57-83.

Bala, N. (1999). Child witnesses in the canadian courts: Recognizing their capacities and their needs. *Psychology, Public Policy and Law, 2,* 323-354.

Bjorklund, D.F., Miller, P.H., Coyle, T.R., & Slawinski, J.L. (1997). Instructing children to use memory strategies: Evidence of utilization deficiencies in memory training strategies. *Developmental Review,17,* 411-441.

Brainerd, C. J., & Renya, V. F. (1990). Gist is the gist: Fuzzy trace theory and new intuitionism. *Developmental Review, 10,* 3-47.

Brown, A.L., & Kane, M.J. (1988). Preschool children can learn to transfer: Learning to learn and learning from example. *Cognitive Psychology, 20*, 493-523.

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

Ceci, S. J. & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. American Psychological Association, Washington, DC.

Connolly, D.A. & Price, H.L. (2006). Children's suggestibility of a repeated event versus a unique event: The effect of degree of association between variable details. *Journal of Experimental Child Psychology*, *93,* 207-223.

DeLoache, J.S. (1987). Rapid change in the symbolic functioning of very young children. *Science*, *238,* 1556-1557.

Drummey, A.B. & Newcombe, N.S. (2002). Developmental changes in source memory. *Developmental Science, 5,* 502-513.

Educational Testing Service (2007). *Graduate Record Examination.* Princeton, NJ

Flavell, J.H. (1970). Developmental studies of mediated memory. In H.W. Reese & L.P. Lipsitt (Eds.), *Advances in child development and behavior* (pp. 181-211). New York: Academic Press.

Foley, M.A., Johnson, M.K. & Raye, C.L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Development, 54, 51-60.*

Giles, J.W., Gopnik, A., & Heyman, G.D. (2002). Source monitoring reduces the suggestibility of preschool children. *Psychological Science, 13,* 288-291.

Gopnik, A., & Graf, P. (1988). Knowing how you know: Young children's ability to identify and remember the sources of their beliefs. *Child Development, 59,* 1366-1371.

Hafstad, G.S., Memon, A. & Logie, R. (2004). Post-identification feedback, confidence and recollections of witnessing conditions in child witnesses. *Applied Cognitive Psychology, 18,* 901-912.

Johnson, M.K., Hashtroudi, S., & Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3-28.

Johnson, M.K., Raye, C.L., Foley, H.J. & Foley, M.A. (1981). Cognitive operations and decision bias in reality monitoring. *American Journal of Psychology, 94,* 37-64.

Kovacs, S. L., & Newcombe, N. S. (2006). Developments in source monitoring: The role of thinking of others. *Journal of Experimental Child Psychology, 93,* 25-44.

La Rooy, D., Pipe, M.E., Murray, J.E. (2005). Reminiscence and hypermnesia in children's eyewitness memory. *Journal of Experimental Child Psychology, 90,* 235-254.

Leichtman, M.D., Morse, M.B., Dixon, A., & Spiegel, R. (2000). Source monitoring and suggestibility: An individual differences approach. In K.P. Roberts & M. Blades (Eds.), *Children's source monitoring* (pp. 257-288). New Jersey: Lawrence Erlbaum Associates.

Lindsay, D.S., Johnson, M.K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology, 52,* 297-318.

Myers, J., Gramzow, E., Ornstein, P.A., Wagner, L., Gordon, B.N., & Baker-Ward, L. (2003). Children's memory of a physical examination: A comparison of recall and recognition assessment protocols. *International Journal of Behavioral Development, 27,* 66-73.

Peterson, C., Pardy, L., Tizzard-Drover, T., & Warren, K. (2005). When initial interviews are delayed a year: Effect on children's 2 year recall. *Law and Human Behavior, 29,* 527-541.

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

Poole, D.A., & Lamb, M.E. (1998). *Investigative interviews with children: A guide for helping professionals.* Washington D.C.: American Psychological Association.

Poole, D.A., & Lindsay, D.S. (1995). Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology, 60,* 129-154.

Poole, D.A., & Lindsay, D.S. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology: Applied, 7,* 27-50.

Poole, D.A., & Lindsay, D.S. (2002). Reducing child witnesses' false reports of misinformation from parents. *Journal of Experimental Child Psychology,81,* 117-140.

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

Pressley, M, Wood, E. & Woloshyn, V. (1990). Elaborative interrogation and facilitation of fact learning: Why having a knowledge base is one thing and using it is quite another. In W. Schneider & F. E. Weinert (Eds.), *Interactions among aptitudes, strategies, and knowledge in cognitive performance.* (pp. 200-221). New York: Springer-Verlag.

Reyna, V. F. & Lloyd, F. (1997). Theories of false memory in children and adults. *Learning and individual differences, 9*, 95-123.

Roberts, K.P. (2000). An overview of theory and research on children's source monitoring. In K.P. Roberts & M. Blades (Eds.), *Children's source monitoring* (pp. 257-288). New Jersey: Lawrence Erlbaum Associates.

Roberts, K.P., & Blades, M. (1995). Children's discriminations of memories for actual and pretend actions in a hiding task. *British Journal of Developmental Psychology, 13,* 321-333.

Roberts, K.P., & Blades, M. (1998). The effects of interacting with events on children's eyewitness memory and source monitoring. *Applied Cognitive Psychology, 12*, 489-503.

Roberts, K.P. (2002). Children's ability to distinguish between memories from multiple sources: Implications for the quality and accuracy of eyewitness statements. Developmental Review. *Special issue on forensic developmental psychology, 22,* 403-435.

Roberts, K.P., & Powell, M.B. (2006). The consistency of false suggestions moderates children's reports of a single instance of a repeated event: Predicting increases and decreases in suggestibility. *Journal of Experimental Child Psychology, 94,* 68-89.

Roebers, C.M. (2002). Confidence judgements in children's and adult's event recall and suggestibility. *Developmental Psychology, 36, 1052-1067.*

Ruffman, T., Rustin, C., Garnham W., & Parkin, A. (2001). Source monitoring and false memories in children: Relation to certainty and executive functioning. *Journal of Experimental Child Psychology, 80,* 95-111.

Thierry, K.L., Lamb, M.E., Orbach, Y. & Pipe, M.E. (2005). Developmental differences in the function and use of anatomical dolls during interviews with alleged sexual abuse victims. *Journal of Consulting and Clinical Psychology, 73,* 1125-1134.

Thierry, K.L., Spence, M.J. & Memon, A. (2000). A comparison between fuzzy trace theory and source-monitoring theory: Evidence from an eyewitness suggestibility study. In K.P. Roberts & M. Blades (Eds.), *Children's source monitoring* (pp. 257-288). New Jersey: Lawrence Erlbaum Associates.

Thierry, K.L., Spence, M.J., & Memon, A. (2001). Before misinformation is encountered: Source monitoring decreases child witness suggestibility. *Journal of Cognition and Development,2,* 1-26.

Thierry, K.L., & Spence, M.J. (2002). Source-monitoring training facilitates preschoolers' eyewitness memory performance. *Developmental Psychology, 38,* 428-437.

Vygotski, L.S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge: Harvard University Press.

Walker, A.G., & Warren, A.R. (1995). The language of the child abuse interview: Asking the questions, understanding the answers. In T. Ney (Eds.), *True and false allegations of child sexual abuse: Assessment and case management.* (pp.153-162). New York: Brunner/Mazel.

Waters, H. S., & Andreassen, C. (1983). Children's use of memory strategies under instruction. In M. Pressley & J. R. Levin (Eds.), *Cognitive strategy research.* (pp. 3-24).

Welch-Ross, M.K., Diecidue, K. & Miller, S.A. (1997). Young children's understanding of conflicting mental representations predicts suggestibility. *Developmental Psychology, 33,* 43-53.

Wellman, H.'M. (1988). The early development of memory strategies. In F.E. Weinert & M. Perlmutter (Eds.), *Memory development: Universal changes and individual differences* (pp. 3-30). Hillsdale, NJ: Erlbaum.

Wheeler, M. A., Stuss, D., T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin, 121,* 331–354.