2010

Training Children Where They Learned Information: A Test of Two Techniques

Justine Renner
Wilfrid Laurier University

Your file Votre référence
ISBN: 978-0-494-68709-3
Our file Notre référence
ISBN: 978-0-494-68709-3

TRAINING CHILDREN WHERE THEY LEARNED INFORMATION:

A TEST OF TWO TECHNIQUES

by

Justine Renner

Bachelor of Arts, Honours Psychology

University of Waterloo, 2008

THESIS

Submitted to the Department of Psychology

In partial fulfillment of the requirements for

Master of Arts

Wilfrid Laurier University

Thursday, September 23, 2010

Abstract

Source-monitoring abilities are crucial skills for children's social and cognitive development, thus, source-monitoring training (SMT) has the potential to benefit children in many practical settings. While some previous research reported that older (7- to 8-years-old) but not younger children (3- to 4-year-olds) benefitted from SMT (Poole & Lindsay, 2002), other studies have found training effects with younger children (Thierry & Spence, 2002; 2004). The current study examined younger and older children's source monitoring trainability by comparing the two different training used in these previous studies: training to a criterion versus a set amount of training. 158 children (aged 3-4 and 7-8) participated in a two- session study. Session one involved watching a DVD and real life demonstration about the body, and a second provided SMT (4-7 days later) about frogs followed by an interview about the previous human body events. Results illustrated that younger and children benefitted from criterion and set SMT, and that training was transferrable to other question types that were not trained. Some benefits of SMT were found for older children, however only for those who were trained to a predetermined criterion. Overall, younger children were found to be less accurate for source information than older children, demonstrating that the ability to monitor source develops gradually over time. However, when children received particularly difficult questions older children in the criterion group, and not younger, more often reported that events occurred in *'real life'*, suggesting that older children may rely on biased reasoning when making difficult source distinctions rather than guessing.

Acknowledgements

This thesis would not have been possible without the guidance and help of several individuals who in one way or another contributed and extended their valuable assistance in preparation and completion of this project.

I would like to thank my advisor, Dr. Kim Roberts, for her continued guidance throughout this project as well as her everlasting support, encouragement, and patience since the day we met. It has been a pleasure to work with her and she truly has been an inspiring mentor to me. I hope we can continue to work together in the future.

Additionally, I would like to thank the members of my committee, Dr. Eileen Wood and Dr. Joanne Lee for their helpful advice, insights, and understanding throughout this project. I would also like to thank Dr. Alan Scorboria for donating his time and reflections for this study.

This research was conducted within the Child Memory Laboratory at Wilfrid Laurier University and could not have been completed without the members, and honorary members of this lab. I would like to thank them for their help in running events, interviewing children, coding, and general data collection, and importantly for their enduring support and encouragement from the start of this project to the end. Each one of these members played a role in the success of this project and I am lucky to have worked closely all of them. Thank you to Suzanne Bell, Sonja Brubachor, Donna Drohan-Jennings, Miriam Fine, Petrice Gentile, Una Glisic, Abigail Goldberg, May Kanipayor, Melissa McCauley, Valerie Patterson, Larissa Pipe, Alexis Rischke, Nasim Shojayi, and Greg Yonick.

Further, I am indebted to the children, parents, teachers, and principals from the local daycares and schools who participated in this research making data collection possible.

Finally, I grateful for my ever patient partner, Julian, my family, and my friends who have endlessly supported my goals and dreams and always believed I would succeed.

Justine Renner

Table of Contents

List of Tables

Training children where they learned information: A test of two techniques

Remembering where one learned information is a crucial skill for cognitive and social development. An inability to do so could have negative effects in educational, social, and forensic settings. Educationally, recognizing source is important for children when deciding the validity of a fact, for example when distinguishing between credible sources (e.g., a teacher) and non-credible sources (e.g., some internet sites). In social settings, remembering conversational partners can avoid embarrassingly repeating a story to the person who originally told it. And in forensic investigations, when cases are built around children's testimonies, it is vital that children can distinguish between sources such as who said what, whether things actually happened or were simply talked about, as well as differentiating between several instances of abuse (Roberts, 2002). Therefore the ability to identify the origin of memories, knowledge, and beliefs, known as source monitoring, is a necessary skill for adaptive functioning (Johnson, Hashtroudi, & Lindsay, 1993).

Further, with increasing technological advances, the knowledge economy is more complex today than ever before. Children who have traditionally been exposed to sources of information from parents, peers, and teachers are now regularly exposed to modern sources such as the internet, television, and increased print media. Thus the crucial ability to monitor source is increasingly multifaceted. It is important to understand how we manage knowledge from these varying and vast sources in order to understand how we monitor source. When we gain information from different sources, what do we do with it and how do we incorporate it into our knowledge? Do we keep source information distinct from other sources or do we incorporate information from one source with another? And because source monitoring is so crucial, is it possible to train children and improve their source-monitoring abilities?

Young children (preschoolers) have difficulty distinguishing between real-life events and television events, and can report events seen in television as though they happened in real-life (Roberts & Blades, 1998). Young children are also susceptible to post-event false information which sometimes results in them describing false information they heard as if it really happened (Poole & Lindsay, 1995). These errors can be detrimental in forensic investigations and are likely the result of children confusing sources, or origins, of events (Poole & Lindsay, 1995). Therefore procedures designed to train children to monitor source would be beneficial. If training was successful, it could potentially prevent mistakes such as these from happening and make children more credible in court. Source-monitoring training (SMT) can also be beneficial to children in educational and social settings as described in the examples above.

However, before training methods can be examined, it is necessary to understand what mechanisms are involved in source monitoring, and its developmental pathway. The next two sections will discuss these topics, in order. Previous SMT research is considered in a third section; closely examining similarities and differences between the few studies in the area. Finally, the last section will introduce the current study.

How do we monitor source?

Unlike straight recognition memory, source monitoring involves both recognition as well as an attribution process in which decisions are made concerning where information was learned. Therefore to monitor source we need to not only recognize that an event occurred, but also make judgments about source. How source judgments are made can be explained by the source-monitoring framework (SMF).

The SMF explains that sources are not given abstract tags or labels during encoding but source judgments are made after an event is experienced while remembering (Johnson et al., 1993; Lindsay and Johnson 2000). According to this framework, there are two main processes responsible for correctly identifying sources of memories. The first process, involves evaluating the qualitative characteristics of memories as they are retrieved. These characteristics include perceptual information (colour and sound), contextual information (spatial and temporal), semantic details (language meaning), affective information (emotions), and cognitive operations (records of elaborating and organizing information), which provide cues to sources. Johnson et al. (1993) claimed to monitor sources one uses the average differences of such characteristics to make distinctions about source. For example, memories high in perceptual, contextual, affective, and semantic information are likely to be memories of externally derived, experienced events. On the other hand, memories that contain more information about cognitive operations (like organizing information read about a place) that took place at the time of the event are likely to be internally derived, non-experienced events. Thus, the more the qualitative characteristics of events overlap, the more likely children are to confuse them (e.g., Lindsay, Johnson, & Kwon, 1991). For example, young children confuse events they saw on television with real-life events because television and real-life are both high in perceptual information and produce similar affective reactions (Roberts & Blades, 1998).

Since source judgments are made by evaluating qualitative characteristics of memories, increasing the similarity among sources makes monitoring sources a much more difficult process. For example, Johnson, Foley, and Leach (1998) asked participants to listen to a set of externally derived words from a confederate and then imagine themselves or the confederate saying another list of words. They found that participants who imagined the words in the confederate's voice performed poorer on source tasks. Lindsay, and Lindsay, Johnson, and Kwon (1990; 1991) found that increasing semantic similarities in two different sources increases source confusions. For example, it is easier to distinguish between two confederates who are talking about different topics (e.g., sports and science), rather than similar ones (e.g., anatomy of different animals). Further, Foley and Ratner (1998) found that if motor actions are similar (such as writing with a pen or imagining writing while holding a pen) more source errors are made than when actions are dissimilar (such as writing with a pen and tracing a line with a finger). These source decisions that are based on qualitative characteristics happen quickly, and theorized to be automatic and unconscious (Johnson et al., 1993).

However, if participants are warned about source confusions or instructed to attend to different sources specifically, fewer source errors are made. This phenomena is explained by the second systematic process in the SMF that in contrast is slower and more strategic and deliberate (Johnson et al. 1993; Lindsay and Johnson 2000). Rather than unconsciously evaluating the qualitative characteristics of memories, this process retrieves supporting information to reason about the source. For example, you may remember seeing a particular painting but know you could not have seen it in a museum you have never been to. So although most source attributions are made during the first more heuristic process (addressing qualitative characteristics in memories), the second systematic process acts as a check for automatic processes and vice versa. In other words, if the threshold for accepting a memory as a perceived event is X, and the amount of perceptual and visual information exceeds X, then it is likely that X was perceived; however, the systematic process may challenge this memory by reasoning it could not have been perceived (like the above painting example). Similarly, the heuristic process can challenge the systematic process if recollections are low on perceptual detail, even if they fit within one's beliefs. For example, the systematic process may result in the decision that you heard a funny joke from a certain friend because it was exactly the kind of joke she would tell; however the automatic process may challenge this on the basis that you do not have enough perceptual details (sound) associate with the memory to determine that you even heard the joke.

While these two processes can help ensure the correctness of source attributions, source monitoring is not an either-or concept (Johnson et al., 1993). It is not unusual to remember the information you learned but not where or when it was learnt. A story that seems familiar can lead you to decide that you must have actually heard it sometime, even though you cannot remember who told you the story and where and when it was heard. Consequently source decisions are made with differing degrees of specificity with varying levels of confidence depending on the type and amount of information available. In general, the ease and accuracy of making a source judgment is dependent upon three things; the type and amount of characteristics for a memory, how unique the characteristics are, and the efficacy of the judgment process.

A second theory, Fuzzy Trace Theory, argues that source errors have nothing to do with the decisions made at the time of retrieval, but are the result of problems with storage and retrieval of the encoded memory traces from an event (Brainerd and Reyna, 1990; Reyna and Brainerd, 1998). According to Brainerd and Reyna, dual representations of each detail of a memory are stored; one representation comprises verbatim details (e.g., what one bought when they went to a store) and the other as more general gist (e.g., what usually happens when one goes to a store). Source is argued to be encoded and represented as verbatim details which are more susceptible to interference from gist, thus memories become more generalized and less detailed as time passes. Poor source-monitoring performance is blamed on the decay of distinct verbatim representations of source. This effect is exaggerated in younger children because they lose verbatim details more easily than older children and adults.

Although Fuzzy Trace Theory is useful to explain how details of general memories may be lost, source-monitoring involves more than just recognition of an event (e.g., *"I saw a stove catch on fire")*. Consisting of both recognizing that X was seen as well as where it was seen (e.g., *"I saw a stove catch on fire while my friend made me dinner"*); source monitoring requires a much finer grain of specificity (Lindsay, 2008). Even if recognition scores in a memory task are almost perfect, source scores may only be at chance (Kahan and Johnson, 1990). While it is reasonable that we store details about source during encoding, when we recall a memory and specify a source, a judgment process must take place to accurately determine where information was learned. It is unlikely for us to know where a specific detail of a topic was learned based on verbatim details alone as Fuzzy Trace Theory claims. Instead it is more likely, as outlined in the SMF, that we can decide where we learned that detail by comparing the qualitative differences of the memory characteristics and using the strategic systematic process to confirm our decision.

Further support for the notion that the SMF better explains the mechanisms behind source-monitoring comes from a study conducted by Lindsay and Johnson (1991) that manipulated source-monitoring tasks to be more difficult or easy. The researchers asked adult participants to study words in two columns either both in depth, or to study one column in depth and the other column briefly. They found that while participants who studied all words in depth performed better on recognition, participants that studied half in depth and half briefly performed better on a source task (distinguishing which list the words originated from). Since Fuzzy Trace Theory argues that a verbatim representation of source is susceptible to quick decay, we would expect that source tasks could not be made easier by instructing participants to study one list more than another. Recognition for words that were studied more deeply may be better than those studied shallowly, but source accuracy would be unaffected, because the verbatim source details would be equal between the two lists. The SMF alternatively explains the increased source accuracy when only one list was studied deeply by the greater qualitative differences in cognitive operations between the two lists. Differences were likely amplified by participants using a memory strategy (like rehearsing the deep list or saying words in that column aloud). Therefore the qualitative differences of the two lists aided by deliberate reasoning (e.g., "I read that word aloud so it must have been in column A") enhanced performance.

In summary, two major theories consider how source monitoring works; SMF and Fuzzy Trace Theory. However, due to the complexity of source monitoring requiring both recognition and a judgment process, it seems that the SMF offers a better perspective into the mechanisms behind making source decisions. Further, since source-monitoring is more complex than mere recognition memory, it is important to understand the development of source monitoring.

Development of Source Monitoring

Generally, children make more source errors than adults until around the age of ten (see Roberts, 2002, for a review). However, in some situations, depending on the complexity of the task, children perform equally as well as adults (see Johnson et al., 1993, for a review). For example, Foley, Johnson, and Raye (1983) had children and adults listen to word lists read aloud by two confederates and later asked them which of the two confederates said specific words. They found that 6-year-olds were as accurate as adults at distinguishing who said what. Similarly, in a study conducted by Roberts and Blades (1998), 4- and 10-year-olds as well as adults watched a real-life event and a video of a similar event and were asked misleading questions about the events a week later. Results showed that although the 4-year-olds were more incorrect, the 10-year-olds performed as well as adults.

Children's source-monitoring ability improves largely between the ages of 3- to 8-years in a gradual, rather than an abrupt, manner (Roberts, 2002). Thus children are better at some types of source monitoring before others (Roberts, 2002). For example, 6-year-olds are better at distinguishing between actions that they performed from ones an experimenter did before they are able to distinguish between actions they actually did versus actions they imagined (Markham, 1991). These differences can be explained by the SMF such that children first learn to use the automatic source-monitoring process before the strategic process. As mentioned earlier, the first process entails examining a memory's qualitative characteristics to determine source, whereas the second process is more deliberate and uses strategic reasoning about source information. However, the more similar the sources are (and thus the more similar the memory characteristics are likely to be), the more likely errors are to occur (Lindsay, Johnson, and Kwon, 1991: Markham et al., 1991). Thus in the example above, the characteristics of remembering doing something versus remembering someone else doing something are more distinct and overlap less than a memory of yourself doing something versus imagining yourself doing something. Lindsay, Johnson, and Kwon (1991) also found 8-year-olds had difficulty differentiating between imagined actions and actions they saw an experimenter perform.

Research using a variety of source tasks also shows a similar pattern of results. Lindsay, Johnson, and Kwon (1991) found a negative correlation between age and source monitoring when 4- and 6 year-olds were asked to distinguish who said what. In this study, the researchers had children listen to two stories read aloud by two confederates and later asked the children who mentioned what specific items. They found that when the speakers were the same gender or when items they were asked about were common to both stories, 4-year-olds performed significantly worse at judging source. Roberts and Blades (1998) discovered that young children also had difficulty differentiating between actions that happened on a video from those that happened in real-life. In this particular study, 4- and 10-year-olds watched a video about a pretend 'children's hospital' and one or two days later watched either a similar (a "birthday party that used very similar items) or dissimilar ("baking a cake that used comparable, but different items) real-life event. One week later children were asked direct questions about items seen (e.g. "Why did the doll cry?"). They found that younger children made more confusions overall in comparison to older children who provided more accurate details. Further, children in the similar condition also made more confusions than children in the dissimilar condition.

Since the study conducted by Roberts and Blades (1998) involved memories for sources that greatly overlap in qualitative characteristics (both have high perceptual and contextual information, with little cognitive operations), children likely had to use the secondary strategic process to reason about learned information. However, as sources become similar on one dimension, other dimensions may become more important to discriminate sources (Johnson et al., 1993: see also Ferguson et al., 1992). Children, especially younger children, likely have problems managing multiple cues or dimensions of sources. Consequently, the children in Roberts and Blades' study would have had to rely on characteristics such as semantic information to differentiate between events due to the fact perceptual and contextual characteristics were so similar, a difficult task for 4-year-olds.

Further, in order to fully evaluate multiple cues and dimensions that are required to source-monitor, children likely need adequate Theory of Mind (ToM) skills which develop around 3- to 4-years-old. A study conducted by Welch-Ross (2000) illustrated those children who passed false-belief ToM tasks paid more attention to a knowledge confederate than an unreliable confederate who provided them with misinformation. As a result, those children reported less misinformation during a later interview. Therefore, ability to reason about mental states can help children manage the cues and dimensions necessary for accurate source-monitoring.

Further, more difficult source problems may require more deliberative and extended strategy use and reasoning that develops later in life (Flavell, 1985). Children may possess a useful strategy and have the mental capability to use it however will not use it unless they are prompted, known as a production deficiency (Flavell, 1970). For example, children as old as 8-years-old fail to organize cards on the basis of meaning, but rather organize them randomly when told to organize them in a way they will remember them (Best and Ornstein, 1986). If they are stressed to organize the cards by meaning, even preschoolers will do so and benefit from it. Children may also suffer from utilization deficiencies such that they possess a strategy, but are unable to benefit from it (Miller, 1990, Miller and Seier, 1994). According to Bjorklund et al. (1997), this utilization deficiency has been demonstrated in 50% of the memory training studies over a 30-year period. Miller et al. (1991) argue that utilization deficiencies occur because strategy use requires numerous mental resources, and children may not have any left to properly use the strategy. It is likely the case that younger children fall within a production or utilization deficiency for strategy use in source monitoring. Children may perform better on source tasks when they are prompted to attend to source, but are unable to do it without that prompting. Younger children could possess the strategies to differentiate between sources, but have a difficult time doing so because correctly using the strategy to come up with a source judgment requires more cognitive capacity than they have available. As mentioned before, source monitoring is more difficult than traditional memory tasks adding extra strategies or even requiring more than one. Recent demonstrations of such source-monitoring strategies have been provided by the work of Ghetti and colleagues. Ghetti and Alexander (2004) found that older children were able to reject misleading information that had high expected memorability (such as going to the Grand Canyon) rather than low memorability (such as eating Corn Flakes for breakfast last week) reasoning that "if I went to the Grand Canyon I would have remembered it". Although training and practice could help children monitor source more consistently, it is often difficult to transfer strategy use to new situations (Cox & Waters, 1986). Thus, with the number of sources and memory characteristics available for events, it can still be difficult for children to develop and rely on source-monitoring strategies without direct instruction.

Younger children's poorer performance on various source tasks can be explained by an inability to use strategic processes to determine source. However, research also indicates that younger children may have an implicit knowledge of source before they are able to strategically reflect on source. Robinson (2000) found that even 3-year-old children rely on informative sources more than uninformative sources. For example, children understood that someone who saw an object pulled out of a box is more informed than someone who never saw it, but only at 5- or 6-years old can children verbally explain why and how they know information (such as what the object is). Younger children on the other hand may report what the object was, but are unable to explicitly state the origin of their knowledge (e.g. "he knows it was a ball, because he opened the box and saw it"). Similarly, Roberts and Blades (1995) found that 3- and 4-year-olds could distinguish memories of performed versus pretended actions when tested nonverbally, but not when tested verbally.

Overall, source-monitoring skills develop gradually throughout childhood. The ease and accuracy of children's source decisions are not solely based on age but also on the extent to which children develop and utilize strategies (as outlined in the SMF), as well as turn implicit knowledge of source into an explicit one. As these tools develop, children become better equipped to make increasingly complex source decisions.

Source-Monitoring Training

Although younger children have more difficulty than older children and adults when monitoring sources, it is possible that young children can perform more similarly to older children if they were given adequate help to drive them toward better source monitoring. As mentioned above, children have some implicit knowledge of source but not the formalized knowledge required to accurately monitor source. Specifically, training techniques could be designed to aid children with implicit knowledge of source develop a more explicit knowledge. In other words, SMT could act as a scaffold for younger children who demonstrate a readiness to monitor source. SMT also has the potential to further children's source monitoring among children who already monitor source helping them distinguish even highly similar sources. Like practicing mathematical problems, practicing source-monitoring skills may further the quality of judgments they can make.

As previously discussed, source monitoring is a crucial skill in social and educational settings, but is especially important during forensic investigations. Young children may agree with interviewers during these investigations because they confuse the origins of events (Poole & Lindsay, 1995; Ackil & Zaragoza, 1995). Similarly, as mentioned before, children have difficulty distinguishing between television and real-life which could result in them incorporating things they saw on television (like ideas of sexual abuse) with their everyday experiences (Roberts & Blades, 1998; Thierry & Spence, 2002). In general, young children who are better at source monitoring are less suggestible than poorer source-monitors (Leichtman et al., 2000; Mazzoni, 1998). Thus training children to think critically about where they learned information could help children more accurately distinguish sources therefore reducing susceptibility to misinformation allowing them to be held admissible in court.

Previous research has examined SMT techniques, however results remain inconclusive. Some researchers have found that only older children benefit from training (e.g., Poole & Lindsay, 2002) while others have found that younger children benefit (e.g., Thierry & Spence, 2002). Two studies, one by Poole and Lindsay, and one by Thierry and Spence are particularly important for the current proposed study.

Poole and Lindsay (2002) examined how training could help children who had been misinformed after an event among three age groups (3-to-4-years-old, 5-to-6-years-old, and 7-to-8-years-old). At a first session, children were shown four science demonstrations performed by an experimenter acting as Mr. Science. Immediately following the demonstrations, children were asked three open-ended questions unrelated to what they had just seen (rapport-building phase) as well as five open-ended questions about Mr. Science without any prompting or reminding. Three months later, parents were asked to read a story about their child's visit with Mr. Science three consecutive days in a row. The stories consisted of four descriptions of demonstrations; one that was explained to the children while they saw it, one that was only seen, one that they simply saw but did not see, and finally one that never happened.

Shortly after hearing the story three times, children visited a different experimenter and were given either SMT or recognition training (control; source was never mentioned) followed by an interview about their visit with Mr. Science. Training consisted of the experimenter performing a series of three "preparation tasks" (like brushing her hair) and talking about other tasks they sometimes do but that the children never saw ("*sometimes I like to add a bit of hair spray to keep my hair in place*"). Children were then asked three open-ended questions about the tasks just performed. Following correct answers (i.e. stating actions they actually saw, not those that were talked about), children were provided feedback (*"that's right, you know that I brushed my hair because you saw me do it").* If children reported an event that they only heard and did not see, they were corrected.

Poole and Lindsay's (2002) interviews consisted of four phases. First, children were asked to freely recall what happened the time they visited Mr. Science, second they were asked open-ended questions about the visit. Third, children were asked direct questions from each of the categories discussed in the story as well as two new instances of touching (e.g. putting a Band-Aid on). In the fourth phase, children in the SMT condition were reminded to only talk about things that actually happened when they saw Mr. Science, referencing the preparation training tasks to explain. Children were then re-asked whether questions they previously answered yes to really happened (e.g. "*Did Mr. Science really put something yucky in your mouth?").* Children in the control group were also re-asked the questions they originally answered yes to, however when reminded to talk about only things that happened when they saw Mr. Science, no reference to the preparation tasks were given.

Not surprisingly, Poole and Lindsay (2002) found that although none of the children reported false events immediately following the demonstration with Mr. Science, false reports greatly increased after the misleading stories were told by parents during the free recall, open-ended questioning, and direct questioning. However, 7- and 8-year-old children in the SMT condition decreased their false reports by half following reminders of source in the fourth phase. This demonstrates that SMT aids older children to draw attention to source relevant information. There were no differences among the younger age groups (3-4-years-old and 5-to-6-years-old) in either SMT or control groups following reminders. Thus Poole and Lindsay concluded that this type of training does not increase younger children's ability to monitor source.

In regards to their results, Poole and Lindsay (2002) offered several explanations why SMT may not have benefitted the younger groups. Most obviously, the three-month period between events with Mr. Science and the story followed by interview could be too long for younger children. Since the same story was repeated three times and occurred shortly before the interview, it is likely that it was confused because children demonstrated a recency effect (remembering most recently presented information better than earlier information; Roberts and Powell, 2007; Baddeley & Hitch, 1977). Along these lines, younger children may not have benefitted because they use less strategic reasoning than older children when making source decisions (Johnson et al., 1993). Thus these children are more likely to 'blurt out' answers to questions rather than think back to the sources (Poole & Lindsay, 2002). Further, stories were read by a trusted adult, likely making it more difficult for children to disagree with misinformation. Moreover, children were not drawn to specific sources, they were told to talk only about things that happened with Mr. Science but were not explicitly told to ignore or distinguish information they heard during the story. Perhaps if children were drawn to specific sources during training (what they actually saw versus what was in the story), SMT could be more helpful. Additionally, SMT may not be enough to help younger children overcome difficultly representing and differentiating between two mental states or representations of events, for instance distinguishing between an actual event versus a description of an event (Templeton & Wilcox, 2000; Welch-Ross, Diecidue, & Miller, 1997). Poole and Lindsay also suggested that 3- and 4-year-olds ability to monitor source may be limited by immature development of the frontal lobe that is important for executive control and intentional retrieval of memory information (2002; Schacter, Kagan, & Leichtman, 1995).

Although Poole and Lindsay (2002) provide an in depth explanation concerning why the 3-to-6-year-olds did not benefit from SMT and the 7-to-8-year-olds did, other research shows that younger children can benefit.

Thierry and Spence (2002) examined 3- and 4-year-olds ability to differentiate between live events and events seen on television after SMT. In their study, children watched six demonstrations performed by Mrs. Science, half were viewed live and half were viewed on the television. Before each demonstration, Mrs. Science clearly labelled the source of the demonstration to differentiate the two sources. Following the demonstrations, a different experimenter took the children into a different room and asked the children a series of eighteen non-misleading questions about the demonstrations (three questions for each demonstration).

Three to four days later children visited with another experimenter and were given either source monitoring or recognition (control) training. Training consisted of watching two puppet shows, one live and another on television. Children in a source monitoring condition were then asked a series of twelve yes/no questions, half of which were misleading. Following any 'yes' response to the questions, participants were asked whether it happened in real-life or on TV. Feedback provided children with the correct source. After each question, children were told they were going to be asked another question but were warned to be careful because some questions are '*tricky*' and are not correct. After the warning, children were asked combination type questions about the question they had just responded to (e.g. *"On the TV, did you see Mrs. Science catch a fish with a magnet?"*), again half of the questions were misleading. Training was complete when children met a criterion of four consecutive correct misleading combination questions and four consecutive correct non-misleading questions. If criterion was not reached, children were asked all twelve questions twice. Children in the control condition were asked the first set of yes/no questions but no mention of source was used. Criterion was met when they answered four consecutive correct misleading yes/no questions and four consecutive correct non-misleading yes/no questions, or they had been asked all twelve questions twice.

Following training, all children were asked a set of 24 questions about Mrs. Science. Twelve of the questions were yes/no recognition questions that were followed by a forced-choice source question for 'yes' answers and the other twelve questions were open-ended questions. Open-ended questions were designed to test transferability of SMT since they were dissimilar to the style of the training questions. Within each type of question, half were misleading. In the misleading yes/no recognition questions, the item was always correct and the source was incorrect and in the misleading open-ended questions, questions were comprised of details that occurred both in the live event and the event seen on television. Responses to yes/no recognition questions were considered correct if both the recognition and source was were correct. Open-ended questions were considered correct if the child correctly described the event or pointed out the false information. Prior to questioning, children were reminded that they could say 'no' to tricky questions and were allowed to say *'I don't know'*.

Thierry and Spence (2002) found that all children met criterion within four to five non-misleading questions and four to five misleading questions. There was also no difference between the two groups when answering the yes/no recognition questions regardless of whether the questions were non-misleading or misleading. However, children in the SMT group were correct more often in response to the source questions that followed the yes/no recognition questions in comparison to the control group who did not differ from chance. In response to the non-misleading open-ended questions, the SMT group made more correct responses while the control group did not differ from chance. However, there was no difference between groups in response to the misleading open-ended questions. Upon further examination, Thierry and Spence broke down errors made during these questions into three categories; between sources (confusing events from one source with another), within events (confusing an item that happened at a different time but from the same source), and confabulations. They found that there were no differences in the within-event confusions and confabulations, but children in the control group made more between-source errors. Thus, SMT did not necessarily benefit children when answering the misleading open-ended questions, but did help avoid making errors between sources.

In addition to showing that younger children can indeed benefit from SMT to a predetermined criterion, Thierry and Spence's (2002) study yielded some additional interesting results. For instance, when asked non-misleading open-ended questions, the SMT group performed much better than the control group who performed at chance. Similar to Poole and Lindsay (2002), Thierry and Spence argued that the control group performed so poorly because they had to consider two representations for one target, a difficult task for young children (2002; Bjorklund, 1995). Thus children have to consider what the interviewer said as well as what they remember about the sources of events. Since source monitoring helps determine how information was acquired, those that were given training answered more questions correctly. SMT also provided children with practice distinguishing between memories with similar characteristics, such as the perceptual and semantic similarities between real-life and television.

As mentioned previously, no difference was found among training groups in the misleading open-ended questions which lead Thierry and Spence (2002) to conclude that SMT was not transferable to other types of questions. Since all the training consisted of yes/no questions and source decisions, children likely had a difficult time transferring those skills to more loosely based open-ended questions. Older children however may have the capability to transfer these skills with the help of a more strategic judgment process, such as retrieving supporting information from the source to correctly describe what occurred, or reject false information (Thierry & Spence, 2002). In other words, the older children did not need as much formal support as the younger children.

Although the results of Poole and Lindsay's (2002) study demonstrated that younger children do not benefit from SMT and Thierry and Spence's (2002) study shows that younger children do benefit, the two studies cannot necessarily be compared with each other. The two studies varied considerably in methodology from the type of sources examined, live versus heard events and live versus televised events, to the type and amount of training used, describing three preparation tasks versus a series of recognition with source questions until criterion is met. Thus it is impossible at this time to conclude whether SMT, is really is beneficial or not.

Current Study

Since SMT has the potential to be used in such practical settings such as education, social, and in forensic investigations to the benefit of children, it is crucial to further test what type of training technique yields the best results. The current study tested both Poole and Lindsay`s (2002) and Thierry and Spence`s (2002) SMT techniques, controlling for all other methodological differences (type of sources used and amount and type of training) in an effort to see at which age and with what type of training children will benefit. The aim of this study was not to replicate the two studies, but to see which of the two training methods (reaching criterion or answering a set number of questions) are better under controlled conditions.

To examine developmental differences in source-monitoring trainability, two age groups were examined, a 3-to-4-year-old group and a 7-to-8-year-old group. Similar to Poole and Lindsay (2002) and Thierry and Spence (2002), children were given science lessons and learned all about the human body through an activity and a DVD. These sources were chosen rather than seen and heard about events because they have more similar memory characteristics such as perceptual information and similar affective reactions, thus making distinctions more difficult.

Thierry and Spence (2002) tested children at short delays (3-4 days), while Poole and Lindsay (2002) used longer delays (3 months). To fairly compare the two types of source-monitoring training in the current study, however, we chose to include a delay longer than in Thierry and Spence's study and shorter than in Poole and Lindsay's study. Specifically, children were tested at a time delay of four to seven days because some forgetting would have occurred at this point, but not so much that the children were merely guessing. Training consisted of first learning about frogs through an activity and a DVD followed by training questions. In order to incorporate the training techniques from both of the previous, three training groups were studied; two SMT groups: a set-training group and a criterion-training group, and a recognition only (control) group. Similar to Thierry and Spence, and Poole and Lindsay, recognition training consisted of asking children a series of four yes/no recognition questions about frogs with feedback provided praising the child or giving them the correct answer. The set-training group was designed to mimic the training procedure used by Poole and Lindsay. Children received the same four questions as the recognition-only group with feedback followed by a forced choice of source to 'yes' answers, again with feedback. The third group, the criterion group, was similar to the training technique used by Thierry and Spence. Training for this group consisted of asking questions in the same format of the set group, however children were asked questions until they reach criterion. Criterion was considered reached when children answered four consecutive non-misleading recognition-source pair questions correctly, or until all 24 questions were asked. However, like Thierry and

Spence found, it was expected that most children (especially older children) would reach criterion relatively quickly with few requiring all 24 questions.

Half of the training questions were based on the real-life demonstration and half about the DVD. Unlike both previous studies, training questions were not equally divided into non-misleading and misleading questions. Rather, a three to one ratio of non-misleading to misleading questions was used. This decision was made because training is meant to be source specific and if yes/no recognition questions are misleading, the correct answer to them is 'no', thus no follow-up source question could be asked.

Immediately following training, children were asked a series of 32 questions about events they saw when they learned about the human body. Sixteen of the 32 questions were yes/no recognition questions followed by a forced choice source question for 'yes' responses, and the remaining 16 questions were recognition-source combination pairs. Four out of 16 questions for each question type were misleading, again, fewer misleading questions will allow us to examine how children source-monitor. These question types were designed such that children were directly asked about two distinct sources unlike Poole and Lindsay's (2002) study. Also, since recognition-source combination questions were not used in training, they acted as a measure of SMT transferability. It is important to note that these recognition-source combination pairs are more difficult questions than the recognition with forced-choice questions because they involve both recognizing an item occurred as well as where it occurred. The misleading recognition-source combination questions are particularly difficult because the item in the question was correct, however the source was not, such that accuracy involved avoiding a *'yes'* answer based on recognition and answering on the basis of source.

As the current study incorporated two techniques used by previous research, a combination of those results was likely. It was expected that the younger children would be less accurate overall in comparison to the older children. Younger children were expected to reach criterion later than older children, and thus younger children in the criterion condition would have had more practice and would make more accurate source responses than both the set and recognition groups of the same age, replicating both Poole and Lindsay's and Thierry and Spence's (2002) results. Since older children were expected to reach training criterion quickly, it was expected that both the set-training and criterion-training groups would provide more accurate source responses in comparison to the recognition only group. This result would parallel Poole and Lindsay's (2002) finding that older children benefit from SMT.

Further, because recognition-source combination questions are more difficult, and were not used as part of the training, it was expected that scores for these *non*-misleading questions would be lower overall. Since older children are better differentiating between sources to begin with and would benefit more from training in general, if training transfer occurred, it was expected to be among older children in the set and criterion conditions. Moreover, because the *misleading* recognition questions are so difficult, no effect of training was expected.

Finally, replicating previous research showing children often confuse and incorporate memories from television for real-life, younger children were expected to be more biased toward saying something occurred in the activity when it actually happened in the DVD than older children. In this exploratory hypothesis, younger children are more likely to make these decisions because they perceive experienced events as more memorable. Thus, if children are unsure of the origin of an event but remember it, they will likely reason that the event must have happened in real-life. However, it was expected that this effect would be smaller for younger children in the criterion-training group as they were expected to benefit from SMT. Similarly, older children were not expected to show any bias toward one source over another, especially for the criterion and set groups.

In summary, younger children were expected to be less accurate overall in comparison to older children. Younger children in the criterion-training group were expected to benefit from SMT whereas older children in the set-training and criterion-training groups were expected to benefit. Additionally, scores for recognition-source combination questions were expected to be lower overall in comparison to yes/no recognition with forced-choice source. Finally, younger children, but not older, were expected to show a bias toward claiming information happened in the DVD.

Methods

Participants

Two hundred and seven 3- to 4 year-olds and 7- to 8-year-olds were recruited from local daycares and elementary schools to participate in this study. Since there are vast development differences in language and theory of mind skills between young 3-year-olds and older 4-year-olds, an emphasis was placed on recruiting 3-year-olds that were 3.5 years or older. A total of 25 children were excluded from analyses. Exclusions were evenly distributed among conditions; however a larger number of younger children (22) were removed then the older children (3). Of the 22 excluded younger children, nine showed a yes or no bias and 16 were either absent for the interview or did not complete it. Three older children were removed because they were absent for the interview. Thus, 158 children aged 3 to 4years old (*N* = *67, M* = 4.45, *SD* = 0.51; Range = 3.44 -5.33) and 7 to 8 years old (*N* = 91, *M* = *7.9, SD* = 0.55; Range = 7.05 – 9.22) were included in the study for data analyses. All children who participated in this study had informed consent from a parent or guardian, and also agreed to participate themselves. Upon completion of the study, all children were rewarded for a job well done and their school received a financial donation for supplies.

Prior to participation, children within each age group were randomly assigned to one of three conditions: either one of two source-training groups (criterion training, set training), or a recognition-training group (control). Hence the final break down of younger participants is as follows: 25 in the criterion-training group, 22 in the set-training group, and 20 in the recognition-training group. Break down for older participants is as follows: 31 in the criterion-training group, 32 in the set-training group, and 28 in the recognition-training group.

Materials and Procedure

Session 1 – The target event. A female researcher informed children that they were going to learn all about the human body and escorted children in groups of four to six into an empty classroom. The human body lesson consisted of a six minute age-appropriate DVD and a similar interactive activity (referred to as the real-life demonstration). These topics were chosen because they are relatively unfamiliar to children and thus children will be more likely to correctly answer questions on the basis of episodic memory rather than general knowledge. Presentation order of the DVD and demonstration was counterbalanced across conditions.

Both the DVD and real-life demonstration consisted of information about the circulatory, respiratory, nervous, and digestive systems; however none of the information provided in the DVD overlapped with the real-life demonstration and vice versa. For example, children may have learned about what the heart looks like in the DVD, but learn how it works in the real-life demonstration. The real-life demonstration involved interactive props and experiments to convey information about the human body (e.g., adding sand to water and pouring it through a filter to show the kidneys filter out toxins). The DVD was similar but featured a different female researcher than the one conducting the live event. Both source modalities were designed to be both educational and enjoyable for the children. See Appendix A for script examples from each source modality.

A set of 16 target items were chosen from each modality based on specific actions and perceptual details which later served as the basis of the target interview (Appendix A for presented items). All items were clearly presented and verbally described to ensure children paid attention to the items. For example, children may have heard it is important to wear a helmet to protect your brain while riding a bike, watched a confederate put on a helmet and state they are ready to ride a bike, and see a picture of a bike. To ensure no confounding of items based on source modality, two scripts were made for each the DVD and real-life demonstration. Thus children were randomly assigned to one of two versions.

Prior to watching the DVD or participating in the real-life demonstration, children were specifically told what was going to happen, providing them with labels for the events (e.g., *"Now that we watched the DVD, we are going to see a real-life* *demonstration"*). The labels DVD and *real-life demonstration* were repeated throughout the target event and training (see below) to ensure children were clear about the labels for each source.

To ensure children did pay attention during the DVD and demonstration, a baseline questionnaire was administered individually to each child immediately after. The questionnaire will consisted of ten questions relating to episodic details of the DVD and demonstration, five from each source. For example, children might have been asked what colour the stethoscope was. Importantly, there was no mention of source in any of the baseline questions.

Session 2. After a 4 to 7 day delay, children were approached individually by another research assistant and asked if they would like to learn about frogs. Once children agreed, they were escorted to an empty room in the school for a follow-up session consisting of a training phase (about frogs) followed by a target interview (about the human body).

Training session. Children learned about frogs through the same source modalities that were used in human body lesson. They watched a four minute DVD as well as participated in four minute real-life demonstration (order was counterbalanced) about frogs that also involved props and interactions. Like the human body lesson, information provided in one modality did not overlap with the other. Two scripts were created and children were randomly assigned to determine whether script one was seen as a DVD or as the real-life demonstration. A set of 24 items for each modality was selected from each script to serve as training questions. Like the human body lesson, all items were clearly presented and verbally described to ensure all children paid attention to them.

Once children watched the frog DVD and real-life demonstration, all props were covered with a blanket and training began. Children in the set- and criterion-training groups received SMT. This training involved asking children a series of recognition-source questions pairs providing feedback on the accuracy of their responses. Specifically, children were first asked a recognition question about something they just witnessed (e.g., *"Was there a blue and orange frog?"*), and following a correct 'yes' response were asked for the source of the information having to choose between the DVD or real-life demonstration (e.g., *"Did you see that in the DVD or real-life demonstration?"*). The order in which the source options were given were counterbalanced. Feedback was given to correct source questions to reiterate the importance of the source (e.g., *"You're right; there was a blue and orange frog in the DVD").* When a child incorrectly responded 'no' to the recognition question, feedback was given (e.g., *"Actually, there was a blue and orange frog"*) and the child was then asked the source question, again with feedback. To ensure children learn to respond 'no' to questions regarding information they did not witness, some misleading (incorrect) questions were asked. Since these items did not occur in either the DVD or real-life demonstration, children were not asked to report the source of the information even if they incorrectly respond 'yes', but were given feedback explaining that the item was not there. See Appendix B for the first four training questions and protocol. Upon completion of training children were given encouragement to think about the sources of the information they learned and also to respond 'no' to things they did not witness (e.g., *"You did a great job, now you know to think about where you learned things from. And you know to say no if something didn't happen").* See Appendix C for specific script following the training session.

Children in the criterion-training condition received training (i.e., given recognition-source pair questions (18) along with a six misleading questions followed by feedback) until they correctly answer four consecutive recognition-source pairs (children must respond correctly to both the recognition and source part of the pair), or until 24 questions have been asked. Note that six of the 24 questions were misleading recognition questions. Children in the set-training received the same source training and feedback however they were only asked three recognition-source pairs and one misleading recognition question. To ensure children did not have a 'yes or no' bias, the misleading question (or first misleading question for the criterion-training group) was always the third question.

Children in the recognition-training condition were asked the same four recognition only questions (e.g., *"Was there a blue and orange frog?"*) with feedback (e.g., *"You're right there was a blue and orange frog"* or in response to an incorrect 'no' response *"Actually there was a blue and orange frog"*). Again, the third question in this set was misleading. Note that no questions about source were asked and children received similar motivation but were not reminded to remember the sources of information (e.g., *"You did a great job, now you know to say no if something didn't happen"*).

Target Interview. Once children completed training, the researcher gave a naïve introduction to the human body lesson (e.g., *"I heard (*researcher*) came in to teach you* *about the human body last week. I wasn't here that day and I don't know anything about it. I'd like you to tell me about the things you learned with* (researcher) *on that day"*). Children were then reminded again to think about where they learned information (criterion- and set-training groups only) and to say no if something did not happen.

Children were asked about each of the 32 items (16 from each source) previously selected from the human body lesson. The types and formats of questions were similar to those used in previous studies (e.g., Thierry & Spence, 2002; Poole & Lindsay, 2002). Half of the items were asked as yes/no recognition questions (e.g., *"Did you listen to someone's heart?")* followed by a forced choice source question (e.g., *"Did that happen in the DVD or real-life demonstration?")*, and the remaining half were asked as yes/no recognition questions combined with source (referred to as recognition-source combination question, *"In the DVD did you listen to someone's heart?").* To ensure children were not simply guessing and did not have a yes bias, eight of the 32 questions were misleading, hence the ratio of non-misleading to misleading questions was 3:1. Since we were interested in the benefits of SMT, it was decided that asking more non-misleading questions would provide more information, given that children were not required to distinguish between sources if something did not happen. Thus, 16 questions were yes/no recognition; 12 non-misleading (six from the DVD and six from the real-life demonstration), and 4 misleading (two about the DVD and two about the real-life demonstration). Similarly, the remaining 16 questions were recognition-source combination questions; 12 non-misleading and 4 misleading with half of each type from the DVD and the other half from the real-life demonstration. Note that all misleading questions consisted of details the child never saw in either source modality. No feedback was provided for the children at any time during the target interview and the researcher was blind to which version of the human body lesson the child was in. See Appendix D for examples of each question type.

Items were counterbalanced in such a way so that every item was asked as a misleading question (either misleading yes/no recognition or misleading recognition-source combination) and as both a non-misleading yes/no recognition question and recognition-source combination question. Thus, four versions of questionnaires for each of the two Human Body versions were created (a total of eight questionnaires) and children were randomly assigned to them. To ensure children did not develop a source bias due to the order in which the modalities were asked, order of modalities were switched in each question. In other words, children may be asked *"Did this happen in the DVD or real-life demonstration"* then the following question will reverse the order asking *"Did this happen in the real-life demonstration or DVD"*. Researchers also asked the set of 32 questions in a random order. This ensured that questions did not follow the same order in which the items were presented and that questions about the two sources were spread out.

Once children finished the target interview, they were praised for a job well done and escorted back to class.

Scoring

Participants were given one point for each correct answer in the baseline questionnaire to determine an accuracy score out of ten. Since children in the criterion-training group received more training, only the first four yes/no recognition scores for training across groups was scored. Similarly, only the first three forced choice source questions were scored for the criterion- and set-training groups. Children were awarded one point for each correct yes/no recognition score and children in the criterion- and set-training groups were given one point for each correct forced choice source. Finally, participants were given one point for each correct yes/no recognition question, one point for each correct forced choice source question, and one point for each correct recognition-source combination question in the target interview. If participants were corrected about incorrect recognition questions and then asked about the source later in the interview and answer source correct, they were given one point.

Results

Analytic Strategy

To better understand any training effects in the target interview, we conducted preliminary analyses on recognition and source accuracy scores before the target session. A baseline recall score measured initial memory for the events, and recognition (yes/no questions) and forced-choice source scores (for the criterion- and set-training groups only) during the training session were used to determine if the training worked.

Measures of recognition, source accuracy, and source biases were obtained at the target interview. Transfer of training was measured by the following dependent variables at the target interview: source accuracy and biases (forced-choice source questions and recognition-source combination questions), recognition accuracy (yes/no recognition questions), correct rejection, and source accuracy in a second attempt (forced-choice source questions to non-misleading recognition questions that were later corrected). See

Table 1 for a complete listing of all variables used in the analyses with descriptions of how they were calculated.

Preliminary Analyses

Baseline questionnaire. The proportion of accurate responses to the baseline questionnaire was calculated by dividing the number of correct responses by the number of items asked (max. 10)[1]. The main purpose of this analysis was to check that there were no condition differences in baseline memory. If there were no effects of condition here but there were effects of training in the target interview, we can rule out the possibility that training effects were an artefact of better initial memory for the events. A 2 (Age: 3-4 years, 7-8 years) x 3 (Training condition: criterion, set, recognition) Analysis of Variance (ANOVA) showed that there were no effects of condition, nor an interaction with age, , Fs ≤ 4.10, ps = ns, η_p^2s $\leq .005$. There was a significant main effect of age, $F(1,152) = 119.70$, $p < .001$, $\eta_p^2 = .441$, because younger children ($M = .34$, $SD = .15$; Min = .0, Max = .7) scored lower on the baseline questionnaire than did older children ($M = .63$, $SD = .17$; Min = .2, Max = 1).

Training. Recognition accuracy in the training session was determined by calculating the proportion of correct yes/no answers out of four. The first four questions were used to allow comparison between all training groups since the set- and recognition-training groups were only given four recognition questions in training whereas the criterion group was given up to 24. A 2(Age) x 3 (Training condition) ANOVA was conducted and a significant Age x Training condition interaction was found, $F(2,152) =$

---

[1] The maximum number of questions is indicated because in some situations one or two of the baseline questions was missed either due to interviewer error or the child had to go back to class.

4.61, $p = .011$, $\eta_p^2 = .057$. Follow-up post-hoc *t-tests* (alpha = .016) revealed that, although younger children performed worse in comparison to older children in all three training conditions (See Table 2), *t*s $\geq - 7.72 \leq -3.390$, *ps* $\leq .001$, η_p^2s $\geq .96 \leq 2.31$, the age difference was particularly pronounced in the recognition condition because the younger children had the lowest scores and the older children had the highest scores.

For the criterion- and set-training groups only, source accuracy for the training session was established by calculating the proportion of correct forced-choice source answers out of three. The first three questions were used so the two groups could be compared, given that the criterion-training groups was given up to 18 forced-choice source questions, whereas the set-training group was only given three. A 2 (Age) x 2 (Training condition) ANOVA determined that there was a main effect of age, $F(1,105) = 21.74$, $p < .001$, $\eta_p^2 = .172$, such that younger children ($M = .64$, $SD = .31$) were less accurate about source than older children ($M = .88$, $SD = .23$). There were no significant effects of training condition or a significant age by condition interaction, Fs ≤ 4.10, ps = ns, η_p^2s $\leq .004$.

Finally, children who met criterion in the criterion-training group (correctly answered four consecutive yes/no recognition and forced choice source questions) were examined. Fewer of the younger children (n=16 out of 25) met criterion whereas all of the older children did (n=30). A 2 (Age) x 2 (Criterion: met, not met) Chi Square test confirmed that fewer younger, but more older, children met criterion than expected by chance, $\chi^2(1, N = 55) = 8.08$, $p = .004$, $\eta_p^2 = .383$. In addition, an independent t-test revealed that younger children also required more trials to reach criterion ($M = 9.32$, *SD*

= .4.72; Min = 5, Max = 21) than did older children (M = 6.33, SD = 2.2; Min = 5, Max = 14), $t(47) = -3.390 \leq 3.00$, $p = .004$, Cohen's $d = 0.9$.

Summary. Younger children were less accurate than older children when responding to baseline memory questions, although this did not interact with training condition. In training, younger children were less accurate than older children when responding to yes/no recognition questions, although this difference was most pronounced in the recognition condition. Finally, the younger children who had received source training (criterion and set groups) were less accurate at identifying source than older children, and also met criterion (criterion group) less often and needed more trials to do so than the older group.

Target Interview

Recognition accuracy. Two recognition accuracy scores were obtained; one for non-misleading yes/no recognition questions, and one for misleading yes/no recognition questions. Non-misleading recognition accuracy was the proportion of correct '*yes*' responses to non-misleading yes/no recognition questions over the total number asked (max. 12)[2]. Similarly, misleading yes/no recognition accuracy was the proportion of correct '*no*' responses to misleading yes/no recognition questions over the total number asked (max. 4).

A 2 (Age) x 3 (Training condition) ANOVA investigating differences for non-misleading recognition questions illustrated a significant effect of age, $F(1,152) = 25.06$, $p < .001$, $\eta_p^2 = .142$, such that the younger group (M = .64, SD = .25) were less accurate

[2] The maximum number of questions for interview questions is indicated because in some situations one or two of the baseline questions was missed either due to interviewer error or the child had to go back to class.

than the older group ($M = .81$, $SD = .16$). Training condition and the interaction were nonsignificant, $F\text{s} \leq .888$, $p\text{s} = \text{ns}$, $\eta_p^2\text{s} \leq .012$. The 2 (Age) x 3 (Training condition) ANOVA examining misleading recognition questions revealed a parallel significant effect of age, $F(1,151) = 6.21$, $p = .014$, $\eta_p^2 = .4$, where the younger group ($M = .45$, $SD = .30$) were less accurate than the older group ($M = .56$, $SD = .27$). Again, neither training condition nor the interaction were significant, $F\text{s} \leq .34$, $p\text{s} = \text{ns}$, $\eta_p^2\text{s} \leq .005$.

Source accuracy. Three types of questions were used to measure source accuracy; recognition-source combination questions (non-misleading and misleading) and forced-choice source questions following a correct response to a non-misleading recognition question. Although the combination questions did require recognition skills they were considered more heavily weighted on source as the source was the part of the question that determined whether or not the correct answer was yes. For example, when children were asked misleading recognition-source questions the item in the question was correct but the source was not.

Non-misleading recognition-source combination questions. The proportion of correct yes responses to non-misleading combination questions over the total number of non-misleading combination questions asked (max. 14) was used in the following analyses. A 2 (Age) x 3 (Training condition) ANOVA revealed a significant main effect of age, $F(1,152) = 19.55$, $p < .001$, $\eta_p^2 = .114$, such that younger children ($M = .67$, $SD = .25$) performed worse than older children ($M = .79$, $SD = .14$). An Age x Training interaction, $F(2,152) = 4.28$, $p = .016$, $\eta_p^2 = .053$ was also revealed. Two one-way ANOVAs with Post Hoc comparisons investigated the effect of training condition within each age group. A main effect of training condition, $F(2, 64) = 3.43$, $p = .039$, $\eta_p^2 =$

.097, was found for the younger children (see Table 3 for means), but not for the older children, $F(2, 88) = 1.99$, $p = .143$, $\eta_p^2 = .043$. Regarding the younger children, those in the recognition group had significantly lower scores from those in the set and criterion groups (*LSD* s $\leq$ .027). However, scores did not differ between younger children in the set and criterion groups (*LSD* s = .896).

Misleading recognition-source combination questions. Misleading recognition-source combination scores were determined by dividing the number of correct *'no'* responses to misleading combination questions by the total number of questions asked (max. 4). A 2 (Age) x 3 (Training condition) ANOVA failed to show any main effects of age or training condition, or interaction, $Fs \leq .34$, ps = ns, $\eta_p^2 s \leq .005$. A series (6) of planned single sample *t-tests* were conducted to test the scores against chance performance (0.5).. Separate *t-tests* were necessary to see which, if any specific training conditions children were performing at chance, and in which age groups this occurred. Younger children in the set group (see Table 4 for means) scored significantly lower than chance, $t(21) = -2.87, p = .009$, Cohen's $d = .96$. Similarly, younger children in the recognition group scored marginally significantly below chance $t(19) = -1.74, p = .098$, Cohen's $d = 1.45$. However, younger children in the criterion group did not differ from chance, $t(25) = -.35$, p = ns, Cohen's $d = 1.45$. A parallel pattern was found with the older children; the set and recognition groups were significantly less accurate than chance, $ts \geq -2.52 \leq -2.91, ps \geq .007 \leq .018$, Cohen's $d \geq .94 \leq 1.18$, but again, the criterion group scored at chance, $t(30) = -1.69$, p = ns, Cohen's $d = 1.24$. Since children in the criterion-training condition were scoring at chance, it could be argued that they were guessing. However, it is more likely to be a result of a training effect since the two groups (criterion and set) who scored at chance were the groups who received the most training, and there were no recognition accuracy differences between any of the training groups within age group. Further, the younger children who scored at chance were the same children who benefitted from SMT as seen in the non-misleading combination questions.

Forced-choice source questions. To correctly answer a forced-choice source question, children had to correctly respond *'yes'* to the corresponding non-misleading yes/no recognition question, thus the number of forced-choice source questions varied by child. The proportional score was found by dividing the number of correct responses to forced-choice source questions by the total number of forced-choice source questions a child was asked (max. 12). Since younger children were less accurate for yes/no recognition, they were asked fewer source questions ($M = 7.54$, $SD = 2.92$) than older children ($M = 8.9$, $SD = 1.9$), $t(156) = -3.59$, $p < .001$, Cohen's $d = 0.58$. A 2 (Age) x 3 (Training condition) ANOVA revealed a main effect of age, $F(1,152) = 16.5$, $p < .001$, $\eta_p^2 = .98$ such that younger children ($M = .65$, $SD = .21$) were less accurate for source than older children ($M = .77$, $SD = .14$). Main effect of condition and the interaction were nonsignificant, $F\text{s} \leq 2.01$, $p\text{s} = \text{ns}$, $\eta_p^2\text{s} \leq .026$.

Summary. Younger children were less accurate than older children when responding to non-misleading source questions; recognition-source combinations and forced-choice source questions. A training effect was observed because younger, but not older, children in the recognition group were less accurate than those in the criterion and set group when answering non-misleading combination questions. Although there were no effects of age or training on responses to misleading recognition-source combination questions, children in the set- and recognition-training groups, regardless of age, performed significantly lower than chance, whereas children in the criterion groups managed to at least reach chance responding.

Source biases. Although children could correctly identify an item from the events, they sometimes erred in their choice of source when asked a follow-up forced choice source question. Since there was not a significant effect of training for these responses, analyses on the children's source errors were conducted to see whether they were biased towards choosing one source over another. Additional source biases were examined for children who incorrectly responded *'yes'* to misleading recognition questions (referred to as *false-alarms*) and therefore still received forced-choice source questions (although obviously neither source was correct).

Source biases following correct recognition. The proportions of correct forced-choice source questions for both the DVD and real-life demonstration were used for analyses. To determine the proportion for DVD, the number of correct DVD responses was divided by the number of questions when the correct source was DVD (max. 6). The same procedure was used to establish the real-life demonstration proportions (i.e., the number of correct real-life demonstration answers divided by the number of questions when the correct source was real-life demonstration (max 6). A 2 (Age) x 3 (Training condition) x 2 (Source correctness: proportion for DVD, proportion for real-life demonstration) ANOVA with repeated measures on the last factor revealed an Age x Source correctness interaction, $F(1,146) = 5.39$, $p = .022$, $\eta_p^2 = .036$. A series (6) of Post Hoc one-sample *t-tests* (alpha = .008) tested the proportion of correct forced-choice source questions for the real-life demonstration against the proportion of correct forced-choice source questions for the DVD for each age and condition group. Six tests were necessary to see which specific training conditions showed biases, and in which age groups this occurred. There were no significant differences for source correctness in any condition for the younger children (see Table 5 for means and test statistics), $ts \geq -1.7 \leq .015$, ps = ns, Cohen's $d \geq 1.4 \leq 1.98$. Significant differences were found for older children in the criterion- , $t(30) = -4.73$, $p < .001$, Cohen's $d = 5.14$, and set-training groups, $t(30) = -3.01$, $p = .005$, Cohen's $d = 3.83$, such that they more often responded *'real life'*. However, older children's source responses in the recognition condition did not differ significantly, $t(27) = .95$, p = ns, Cohen's $d = 4.11$.

Source biases following false-alarms. Two proportional scores were calculated to examine the false-alarm source question biases; one for each source. For the DVD false-alarm proportion, the number of times children said *'DVD'* in response to a false-alarm forced-choice source question was divided by the number of false alarms. To examine whether age or training condition affected false alarm source, a 2 (Age) x 3 (Training condition) x 2 (False-alarm source: proportion DVD, proportion for real-life demonstration) ANOVA with repeated measures on the last factor was conducted. There was a significant Age x False-alarm source interaction, $F(1,152) = 6.32$, $p = .013$, $\eta_p^2 = .04$. As above, a series (6) of follow-up one-sample *t-tests* (alpha = .008) tested the proportion of time real-life demonstration was mentioned in response to a false-alarm forced-choice source questions against the proportion for DVD. Younger children (see Table 6 for means and test statistics) did not differ regardless of what condition they were in $ts \geq -2.26 \leq .068$, ps = ns, Cohen's $d \geq 1.28 \leq 2.12$. Significant differences were found for older children in the criterion- , $t(30) = 6.34$, $p < .001$, Cohen's $d = 3.47$, and set-training groups, $t(31) = -3.92$, $p < .001$, Cohen's $d = 2.83$, group such that they frequently answered *'real life'* in response to a false-alarm forced-choice source question. There was no significant difference, however, for older children in the recognition-training condition, $t(27) = 1.02$, $p =$ ns, Cohen's $d = 3.16$.

Summary. Younger children's source errors in response to forced-choice source questions did not favour one source more often than another, regardless of the training condition. However, older children in the criterion- and set-training groups displayed a bias to say *'real life'* more frequently than *'DVD'*. Older children's response in the recognition-training group did not differ. Similarly, in regard to false-alarm source questions, there were no significant differences in source answers for younger children, however older children in the criterion and set groups were biased to say *'real life'* more than *'DVD'*. Again, older children in the recognition group did not show a preference for one source.

Discussion

The current study sought to examine younger and older children's trainability for source monitoring by comparing the two different types of training used in these previous studies: a set amount of training versus training to a criterion. Source-monitoring abilities are crucial skills for children in many respects including education, socially, and in forensic investigations. Thus, SMT has the potential to benefit children in many practical settings. While some previous research reported that older (7- to 8-years-old) but not younger children (i.e., 3- to 4-year-olds) benefitted from SMT (Poole & Lindsay, 2002), other studies have found training effects with younger children (Thierry & Spence, 2002; 2004).

Consistent with Poole and Lindsay's (2002) results and many other researchers (Lindsay, Johnson, & Kwon, 1991; Roberts, 2002; Roberts & Blades, 1998; Robinson, 2000), clear age differences were found in children's source-monitoring ability. Younger children were consistently less accurate in source and recognition measures than older children. Consistent with Thierry and Spence's (2001) results, training effects were found for the younger children, and the effects in the current study were so robust the training transferred to difficult question types that were not used in training.

Although a set amount of training, or training to a predetermined criterion benefited younger children, the most improvement in each age group was seen after children had been trained until they met a predefined criteria used to signify that they had had enough training. In situations where source-monitoring was not improved by SMT, older, but not younger, children demonstrated a bias towards claiming events happened in *'real life'* more than *'DVD'*. This could suggest that older children may run into difficulty source-monitoring because they rely on the strategy *'if I remember it, I must have seen it'* when they are unsure of source. In this section results will be discussed in depth and conclusions about the efficacy of SMT and source-monitoring development will be drawn. Finally, implications and future directions for this research will be considered.

Evidence of training effects

To determine the effectiveness of SMT in this study, children's source accuracy was measured using two different question types; recognition-source combination and forced-choice source questions. The first measured the transferability of SMT to different question types and the second directly measured the effect of SMT as this format mirrored the one used in training. To further understand how training affected children's ability to monitor source accurately, tendencies toward one source over another were inspected.

Transfer of training. As mentioned previously, recognition-source combination questions were used to measure whether SMT could be transferred to other question types. This was possible because children never saw questions like these during the training phase. These questions were thought to be more difficult than the direct forced-choice source questions because children must simultaneously consider both whether an event occurred and where it occurred. Therefore, it was hypothesized that scores for these questions would be lower overall, and that if a training transfer occurred, it would be for older children in the criterion- and set-training groups. Although not identical, training effects in responses to both the non-misleading and misleading recognition-source combination questions were observed.

Examination of age differences for *non*-misleading recognition-source combination accuracy revealed that, as expected, younger children were less accurate than older children. Older children are more accurate source-monitors in general (Roberts, 2002) and since these questions do require both recognition and source-monitoring all at once, it was not surprising that younger children were less accurate. Although it was predicted that the older children would benefit most from training, given their advanced cognitive control and previously reported cognitive improvements (Poole & Lindsay, 2002), training effects were observed only in the younger age group. Younger children in the criterion- and set-training groups were much more accurate (71% and 72%, respectively, accurate) in comparison to the recognition training group (55%

accurate). In other words, only the younger children demonstrated training transferability.

Analysis of the means from the older children in all three training conditions may give some clues as to why there was no training effect for the older children. Their scores were higher than the younger children, ranged from 76% to 82%, and did not differ from each other, contrary to the hypothesis. Even the recognition group, who did not receive any source training, was 80% accurate suggesting that older children were stronger in general than anticipated in these types of questions. Plausibly, simply having a source mentioned in conjunction with a recognition question was enough to direct the older children to think about the question as a whole more thoroughly. However, older children's non-misleading recognition accuracy scores (ranging between 79-81%) were comparable (i.e., questions probing details with no mention of source). It is possible that older children as a whole said *'yes'* more often but this is unlikely since older children did say *'no'* to some of each question type and were more accurate in general compared to younger children (except in misleading recognition-source combination questions see below). Therefore, the lack of training effect for older children in non-misleading recognition-source combination question is probably because they were generally so accurate and did not need training; there was little room for SMT to make a difference in their scores.

An encouraging result was that both criterion and set training were equally beneficial to younger children. Training was so beneficial, trained younger children's accuracy scores were much closer to the older children's than the recognition group. This replicates Thierry and Spence's (2001) findings that if younger children are trained to a criterion, they can benefit from SMT. On the other hand, these results disagree with Poole and Lindsay's (2002) findings that argue younger children who receive a set amount of training do not benefit from SMT. Additionally, these data show that SMT is transferrable to other question types which were not seen in Thierry and Spence's training study. Transferability of source training was likely significant in this study because the recognition-source combination questions children did not have in training involved the same aspects of the training questions they did practice (recognition with forced-choice source), just put together in one question. In other words, during training children had practice answering recognition questions followed by a forced-choice source question (e.g., *'Did you learn some people eat frogs?'* followed by *'Was that in the DVD or real-life demonstration?'*), so later when the recognition and source components were put into one question (e.g., *In the DVD, was there toothpaste*), children were already familiar with the two different parts of the question. Whereas Thierry and Spence used recognition-source combination and recognition with source questions in training but assessed SMT transfer by asking completely different open-ended questions. Open-ended questions are also much more difficult because they involve recalling memory to construct an answer, where recognition and forced-choice source simply involve picking one out of two options. This suggests that in order for younger children to transfer learning from SMT to other question types, training must consist of similar skills that will be necessary for those other questions.

Misleading recognition-source questions were even more difficult than their non-misleading counter parts. In the non-misleading questions, both the item and source were correct whereas only the item in the misleading question was correct - the source was wrong. To answer correctly then, children had to recognize that the item occurred, but realize that it did not happen in the mentioned source. It was not unexpected that there were no age differences in accuracy for these questions, or that no training effects were found. Both younger and older children's score were considerably low (accuracy between 30-47% for younger and 35-40% for older) across condition. SMT may not have been effective simply because these questions were too difficult and the training children received may not have been enough to successfully answer these questions.

Further examination into individual training conditions within each age group illustrated some interesting patterns. Accuracy scores for misleading recognition-source combination questions were tested against chance performance in an effort to see if children were randomly guessing answers. Younger and older children's scores in the criterion-training group did not differ from chance. However all children in the set- and recognition-training group had scores significantly lower than chance. These scores suggest that although there was not a clear significant effect of training for these difficult questions, training was advantageous in some ways. Children who were trained until they met a training criterion (or were given a large amount of training) scored at chance, thus, an appropriate amount of source training can improve children's abilities (in a small way) to transfer training to very difficult questions. Hence, these results further suggest that younger children can benefit from training and that training can be transferred. However, they also imply that to be successful at more difficult questions that they did not receive training for, both younger and older children may require more practice. Similarly, children may also answer the recognition questions and simply do not know how to dig deep enough to reflect on source unless they have had some practice.

Forced-choice source questions. Forced-choice source questions were used in source training and therefore were a direct measure of SMT. Accordingly then, it was hypothesized that the most training effects would be found for these questions; specifically that younger children in the criterion group would benefit (replicating Thierry and Spence's results, 2002) and older children in both the criterion and set group would benefit (replicating Poole and Lindsay's results that set training was beneficial, 2002). Furthermore, it was expected that younger children would score lower overall compared to older children. While the younger group did score lower than the older group, no effects of training were found for forced-choice source questions, and hence did not directly replicate the previous results of Thierry and Spence (2002) or Poole and Lindsay (2002).

Younger children's accuracy ranged from 62-68%, and older children's ranged from 74-81%. Plausibly then, children (especially younger children) may have tended to choose one source more than another when asked whether the item occurred in *'Real life'* or *'DVD'*. A further possibility was that older children were already fairly accurate and there may have been little room for SMT to be beneficial. To investigate these predictions additional analyses examining children's biases toward source were conducted.

Source biases. Source biases were in examined in two ways; first biases for forced-choice source after a correct *'yes'* to a non-misleading recognition question were examined, followed by biases when children incorrectly responded *'yes'* to a misleading recognition question (false alarms). Within both of these analyses it was hypothesized that if a source bias occurred, it would occur in younger children such that they would have the tendency to answer *'Real life'* more than *'DVD'*. This would suggest that when younger children recognize an item but source memory is vague assume it must have happened in real life because actually experienced events (i.e., Real life) are more memorable.

Source biases following correct recognition. Forced-choice source biases were determined by comparing children's source accuracy for DVD with their accuracy for Real life. Unexpectedly, older children in the criterion- and set-training groups demonstrated a bias for Real life. Older children in the recognition-training group and all of the younger children did not have a preference for source as their answers did not differ from chance. In other words, the older children who received SMT were the ones who showed the greatest bias. Plausibly then, source training may have caused the older children to reflect about source when answering forced-choice source questions. Although they were not any more accurate than the other groups, they were trying to use the training practice to reflect upon where items occurred. Thus, when older children did recognize an event but were unsure of where it occurred, they relied on the strategy *'if I remember it, then I must have experienced it in real life'*, because they perceive memories for events experienced in real life as more memorable than those viewed on DVD. Alternately, children may have thought the DVD was more interesting and could have developed a *'I would have remembered it if it happened in the DVD'* bias. However, the later explanation seems doubtful because the Real life demonstration involved many interactive parts with novel props it appears unlikely children would have found the DVD more interesting. Either way, it could suggest that the older recognition group's lack of source practice caused them to simply guess source answers for vague memories.

Source biases following false alarms. False alarm biases were determined by comparing the proportion of times children chose *'DVD'* with the proportion of times they said *'Real life' (when neither source was correct).* Again, older children in the criterion- and set-training group demonstrated a bias to choose *'Real life'* than *'DVD'* whereas older children in the recognition group and all of the younger children regardless of training did not show a bias (scores were at chance). Given that these source choices followed false alarms, 'memory' for these details was probably more vague than those for details that were present. This result furthers the argument that older children who are trained to monitor source but have vague memories use a strategy that experienced memories are more memorable than those viewed from a DVD.

Training. To fully determine why the training was effective for the younger children, it is important to understand how and when children were considered trained within each condition. Children underwent one of three types of training; recognition and source training until a predetermined criterion was met, a set amount of training, or recognition training only. Although younger children were less accurate in response to recognition questions compared to older children (a well known finding: Koriat et al., 2001; List, 1986; Ornstein, Gordon, & Baker-Ward, 1992; Peterson, 1999; Poole & White, 1991; 1993), overall training condition did not relate to recognition accuracy. This was not surprising as younger children often have more difficulty with memory tasks in general, and no training group differences were expected as all three groups received the same first four recognition questions.

Within the groups that were given source training (set and criterion), the first three forced-choice source questions were studied. Again only the questions that both groups received were considered to ensure direct comparisons between the training groups could be made. As expected, younger children were found to be less accurate for source during training than older children, and no differences for training group were found. Therefore by the third forced-choice source question, children in the criterion and set groups could be considered evenly matched within their age group. This means that the training effects in the target interview can reliably be attributed to the criterion training because the set group had completed training after these three questions, but the criterion group continued training until they met criterion or until they had been given much more training.

The lack of a difference in source monitoring between the set and criterion conditions in the older group is easy to explain based on previous research. Because older children met criterion so quickly (needing an average of 6.33 questions), they did not receive much more training than their counterparts in the set condition. Similarities in source monitoring by the younger children in the set and criterion conditions are not as easily explained, however. Younger children in the criterion condition required an average of 9.32 questions to meet criterion, which was more than the older group. Further, fewer younger children met criterion than expected by chance (only 64%), but all of the older children, and more than expected by chance, successfully met criterion. Since younger children did require more trials and not all met criterion, generally the older criterion group received more training than the younger set group. Perhaps then, it was not necessarily the amount of source training that younger children received as much as it was that they received any training at all. Maybe children simply needed to be reminded/alerted to think about where they saw information to demonstrate any source benefits.

Source-monitoring development

As outlined by the SMF (Johnson et al., 1993), source-monitoring is a judgment process that occurs after an event is experienced while remembering. Source-monitoring relies on both making automatic decisions based on the qualitative characteristics of memory and by using a slower more deliberate process to reason about source. When two sources are similar in qualitative characteristics, as the case in this study, one must use the more strategic and deliberate process to make source judgments. This process may be more or less difficult depending on a child's age or the complexity of the questions, however warning children to attend to the source (i.e., providing them with training) can make source judgments more accurate (Lindsay & Johnson, 2000) as replicated in younger children in this study.

This study replicated previous research that the ability to source monitor develops gradually between the ages of 3- to 8-years (Poole & Lindsay, 2002; Roberts, 2002; Roberts & Blades, 1998; Robinson, 2000) since younger children were consistently less accurate than older children during source decisions, and some types of source questions were easier than others. These results are explained by the SMF such that the more two sources overlap in qualitative differences, more importance is placed on using a second strategic process to reason about source (instead of making source decisions based on the first automatic process; Johnson et al., 1993). Further, since the events in this study greatly overlapped in qualitative characteristics (both were high in perceptual and contextual information); other dimensions may have become more important to discriminate sources (e.g., relying on semantic differences between the two sources; Johnson et al., 1993; Ferguson et al., 1992). Younger children likely have problems managing multiple cues or dimensions of sources making source distinctions even more difficult. In order to fully evaluate multiple cues and dimensions required to source-monitor, it seems children need adequate Theory of Mind skills, which develop between starting at around 3-years-old. This can be seen in Welch-Ross's (2000) research were children who have can pass false-belief tasks also remember information from credible sources more than information from an incredible source.

Since source-monitoring requires strategic and deliberate reasoning, SMT gave younger children the practice they needed to become more accurate source monitors. Further, older children in the criterion-training condition used the SMT to develop a strategy for answering source questions for vague memories. This effect was seen in the Real life bias for forced-choice source questions for correct recognition and false alarms. Although these children seemed to possess a strategy of *'if I remembered it, it must have happened in Real life'*, they did not benefit from it because they possessed a utilization deficiency (Miller, 1990; Miller & Seier, 1994). In other words, older children had a strategy to monitor source, but were unable to use it because the strategy may have been too mentally taxing to use correctly after the recognition question for a vague memory.

Although younger children had more difficulty using deliberate source judgments than older children, younger children's improved accuracy as a result of SMT suggests that children who did not receive training may have possessed an implicit knowledge of source. The success of the training with these young children suggests that they were cognitively 'ready' to benefit from instruction and practice in using an advanced skill. As previous research (Robinson, 2001; Roberts & Blades, 1995) found, younger children can distinguish sources nonverbally before verbally showing that although younger children may not be able to tell you where something occurred, does not mean they do not possess the source knowledge. Given that younger children who were given SMT were more accurate, it is plausible that the younger children in the recognition group understood where items occurred but had difficulty explicitly relaying that information.

Further, there were no age differences in children's accuracy for the difficult misleading recognition-source combination questions but older children were much better when these questions were non-misleading. This confirms that because source-monitoring skills develop gradually over time, children are better at some types of source monitoring before others (Roberts, 2002). Thus, it seems that children are better at answering non-misleading source questions before answering misleading questions. This is not surprising since the misleading recognition-source combination questions involved an experienced event but an incorrect source, so to be successful children had to consider their belief that the event occurred, but realize it did not occur in the mentioned source. Hence, these difficult questions required children to use a more deliberative and extended reasoning (e.g., *'I know that I saw something similar to that, but I don't think it was that, and it couldn't have been in the DVD because I remember touching toothpaste right after I learned about it;* Flavell, 1985) that they did not yet possess.

Although the results of this study strongly fit within the SMF, an alternate theory, Fuzzy Trace Theory, could also explain the results. According to Fuzzy Trace Theory, source information is simply another verbatim detail that is encoded at the time of the event and as such is subject to quick decay (Brainerd and Reyna; 1990). Poor source-monitoring then is blamed on the decay of distinct verbatim representations of source. According to fuzzy-trace theory, then, children are poorer source monitors because they lose verbatim details faster than adults. Source training effects then would be argued to be the result of training children to better reinstate the verbatim detail of source. The results of this study argue against this explanation, however, because the two sources were mentioned constantly throughout the interview – children did not need to remember source. Rather, source-monitoring involves more than just memories of verbatim details. Source-monitoring requires both a recognition of an event as well as where the event occurred which requires a much finer grain of specificity (Lindsay, 2008). Importantly, recalling a memory and specifying the source of that memory requires a judgment process to accurately determine where it was learned. Rather than simply judging where an item occurred based on verbatim details alone as Fuzzy Trace Theory argues, it is more likely that we decide the source of information by evaluating the different characteristics of the memory and using deliberative reasoning, as those outlined in the SMF, to confirm our decisions. Benefits of training can then be explained by the SMF such that SMT provided children with the practice necessary to utilize a more deliberative process, a process that young children would not spontaneously use although they have some tacit ability to monitor source. Therefore, the results of this study clearly argue in favour of the SMF, as it explains the strategic reasoning that occurs when source decisions are made (as seen in the biases for vague memories) as well as how training can benefit younger children by providing them with practice to better utilize this strategic process.

Implications and future directions

Because source-monitoring is such a crucial skill for cognitive and social development, this study has many theoretical and practical applications. Specifically, SMT can be beneficial to children in educational, social, and forensic settings. Educationally, recognizing source is important for children when determining the validity of a fact. In this information age children are exposed enormous amounts of information from a variety of sources, for example children may learn about the solar system from a teacher, a book, and several websites. When they sit down to complete a project on the solar system however, it is important that they are able to distinguish between credible and non-credible sources of information. For example, they may remember that Mars is coming close to Earth and that could cause people to have good luck, but it is important that they remember that this information came from a website and not their teacher. Therefore training children to think about where information came from after they learned it could aid them within several educational facets from writing reports to taking tests. Similarly, source-monitoring is important in social settings. During conversations, it is important to remember that the story you are telling was already told to you by the very person you are talking to in order to avoid embarrassment. Hence before you start telling the story, it would be useful to think first about who told you the story originally.

While source-monitoring is important for educational and social settings, perhaps where it is most crucial is in forensic settings. During forensic investigations, it is vital that children are able to distinguish between who said what, whether events actually happened for were seen on television, as well as differentiating between several instances of abuse (Roberts, 2002). Since these investigations can be built entirely around a child's testimony an inability to monitor source could have detrimental effects. This study is encouraging because it shows that children as young as 3-years-old can be trained to be more accurate source-monitors, and thus it is possible that training techniques could be used in the field to help investigators gather more reliable testimonies. However, this study also showed that even when children meet a training criterion, they may develop harmful strategies biasing them toward one source more than another. It is very worrisome that children who received the most amount of training in this study were biased to claim that the item occurred in Real life when memories were vague. In investigations this could lead to false allegations and even convictions.

As a result, future research could look more in depth at the amount and type of source training that is most beneficial to forensic investigations. Perhaps more practice saying '*no*' or '*I don't know*' could help rid children of these biases. Children were given practice saying '*no*' and were told that it was okay to say '*I don't know*' however, source accuracy was the focus of this research and therefore children received fewer misleading questions. Similarly deeper examination into the development of source biases may reveal that at different ages, children have different biases. They may have been biased towards Real life at 8-years-old, however they may reverse their strategy thinking '*I would have remembered it if I saw it in Real life, so it must have happened in the DVD*'.

Further research investigating the transferability of SMT may also be beneficial to forensic investigators. This study found training transfer effects for questions that required similar processing as the training questions, however other literature (Poole & Lindsay, 2002) found no transfer effect when the open-ended style of questions used in forensic investigations were asked. Plausibly then additional training techniques could investigate how children may be trained to better discuss source spontaneously or answer open-ended source questions more accurately.

Finally, an understanding of how long benefits of training last could offer interesting insight further into source-monitoring development. If children are cognitively ready for SMT, it may be possible to train them to be as accurate as adults earlier than expected (10-years-old; see Roberts, 2002, for a review). Additional training techniques could also examine whether even the most difficult source distinctions could be made easier.

Conclusion

In conclusion, this study found that both younger children can be trained to become more accurate source-monitors and older children who receive training to a predetermined criterion may show some benefits as well. It replicated previous research showing that younger children can benefit from training to a predefined criterion (Thierry & Spence, 2001), and that older children may benefit from training (Poole & Lindsay, 2002), although potential benefits were only shown when older children had been trained to a predetermined criterion. It also opposed previous research (Poole & Lindsay, 2002) showing that younger children too can benefit from a set amount of SMT. Further, this research demonstrated clear training transfer effects for more difficult questions not used in training, a result that was not previously found (Thierry & Spence, 2001). Although children benefitted from any amount of training, children who received training until they met predetermined criteria showed the most improvement. Moreover in circumstances where SMT did not improve accuracy, older children, but not younger, demonstrated a bias towards choosing *'Real life'* more than *'DVD'*.

The results in this study are consistent with the SMF (Johnson et al., 1993) and previous research examining the development of source-monitoring (Poole & Lindsay, 2002; Roberts, 2002; Roberts & Blades, 1998; Robinson, 2000). In nearly every situation younger children were less accurate than older children illustrating that the ability to source-monitor develops gradually. A situation where older children were as inaccurate as younger children confirmed the argument that some types of source monitoring develop before others (Roberts, 2002).

Finally, this study has many theoretical and practical applications that can be applied to children's everyday lives (especially educationally and socially), as well as in forensic investigations. While the study answered many questions about the feasibility of training children to better remember the sources of their information, it also leaves many potential avenues for future research to explore.

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. doi:10.1006/jecp.1995.1031

Best, D. L. & Ornstein, P. A. (1986). Children's generation and communication of mnemonic organizational strategies. *Developmental Psychology,* 22. doi:10.1037/0012-1649.22.6.845

Bjorklund, D. F., & Douglas, R. N. (1997). The development of memory strategies. In N. Cowan (Ed.), *The development of memory in childhood* (pp. 201-246). Hove, England: Erlbaum.

Bjorklund, D. F. (1995). *Children's thinking: Developmental function and individual differences* (2nd ed.). CA: Thomson Brooks/Cole Publishing Co.

Brainerd, C. J. & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review,* 10, 3-47. doi:10.1016/0273-2297(90)90003-M

Cox, D. & Waters, H. S. (1986). Sex differences in the use of organization strategies: A developmental analysis. *Journal of Experimental Child Psychology,* 41, 18-37. doi:10.1016/0022-0965(86)90048-2

Ferguson, S. A., Hashtroudi, S. & Johnson, M. K. (1992). Age differences in using source-relevant cues. *Psychology & Aging,* 7, 443-452. doi: 10.1037/0882-7974.7.3.443

Flavell, J. (1970). Developmental studies of mediated memory. In H. Reese & L. Lipsitt (Eds.), *Advances in child e development and childe behaviour* (Vol. 5, pp. 181-811) New York: Academic Press.

Flavell J., (1985). *Cognitive development* (2ed.), Englewood Cliffs, NJ: Prentice Hall.

Foley, M. A. & Ratner, H. H. (1998). Distinguishing between memories for thoughts and deeds: The role of prospective processing in children's source monitoring. *British Journal of Developmental Psychology,* 16, 465-484.

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. Retrieved from http://www.jstor.org/stable/1129860

Ghetti, S., & Alexander, K. W. (2004). "If it happened, I would remember it": Strategic use of events memorability in rejection of false autobiographical events. *Child Development,* 75(2), 542-561. Doi: 10.1111/j.1467-8624.2004.00692.x

Johnson, M. K., Hashtroudi, S. & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin,* 114, 3-28. doi: 10.1037/0033-2909.114.1.3

Johnson, M. K., Foley, M. A., & Leach, K. (1988). The consequences for memory of imaging in another person's voice. *Memory & Cognition,* 16, 337-342. Retrieved from http://memlab0.eng.yale.edu/PDFs/1988_Johnson_Foley_Leach_MemCog.pdf

Kahan, T. L., & Johnson, M. K. (1992) Self effects in memory for person information. *Social Cognition,* 10(1), 30-50.

Koriat, A., Goldsmith, M., Schnieder, W., & Nakash-Dura (2002). The credibility of children's testimony: Can children control the accuracy of their memory reports?

Journal of Experimental Child Psychology, 79, 405-437. doi:10.1006/jecp.2000.2612

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). Mahwah, NJ: Erlbaum.

Lindsay, D. S. (1990). Misleading suggestions can impair eyewitnesses' ability to remember event details. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* 16, 1077-1083. doi: 10.1037/0278-7393.16.6.1077

Lindsay, D. S. & Johnson, M. K. (1991). Source monitoring and recognition memory. *Bulletin of the Psychonomic Society,* 29, 203-205. Retrieved from Lindsay_Johnson_BulPsychSoc.pdf

Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Recognition memory and source monitoring. *Bulletin of the Psychonomic Society,* 29, 203-205. Retrieved from http://memlab0.eng.yale.edu/PDFs/1991_Lindsay_Johnson_BulPsychSoc.pdf

Lindsay, D. S. & Johnson, M. K. (2000). False memories and the source monitoring framework: Reply to Reyna and Lloyd (1997). *Learning & Individual Differences,* 12, 145-161. doi:10.1016/S1041-6080(01)00035-8

List, J. A. (1986). Age and schematic differences in the reliability of eyewitness testimony.

Developmental Psychology, 22, 50–57. doi: 0012-I649/86/S00.75

Markham, R. (1991). Development of reality monitoring for preformed and imagined actions. *Perceptual and Motor Skills,* 72, 1347-1354.

Mazzoni, G. (1998). Memory suggestibility and metacognition in child eyewitness testimony: The roles of source monitoring and self-efficacy. *European Journal of Psychology of Education,* 13, 43-60. doi: 10.1007/BF03172812

Miller, P. H., & Seier, W. L. (1994). Strategy utilization deficiencies in children: When, where, and why. In H. W. Reese (Ed.), Advances in child development and behavior (pp. 107-156). San Diego: Academic Press, Inc.

Ornstein, P. A., Gordon, B. N., & Baker-Ward, L. (1992). Children's memory for salient events:

Implications for testimony. In M. Howe, C. Brainerd, & V. Reyna (Eds.), *Developmental long term retention* (pp. 31–50). Hillsdale, NJ: Erlbaum.

Peterson, C. (1999). Children's memory for medical emergencies: 2 years later. *Developmental*

Psychology, 35, 1493–1506. doi: 10.1037/0012-1649.35.6.1493

Poole, D. A. & Lindsay, D. S. (2002). Reducing child witnesses' false reports of misinformation from parents. *Journal of Experimental Child Psychology,* 81, 117-140. doi:10.1006/jecp.2001.2648

Poole, D. A. & Lindsay, D. S. (1995). Interviewing preschooler: Effects of nonsuggestive techniques, parental coaching, and leading questions of reports of nonexperienced events. *Journal of Experimental Child Psychology,* 60, 129-154. doi:10.1006/jecp.1995.1035

Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*, 27, 975–986. doi: 10.1037/0012-1649.27.6.975

Poole, D. A., & White, L. T. (1993). Two years later: Effect of question repetition and retention interval on the eyewitness testimony of children and adults. *Developmental Psychology*, 29(5),844–853. doi: 10.1037/0012-1649.29.5.844

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. doi: 10.1002/(SICI)1099-0720(199810)12:5<489::AID-ACP535>3.0.CO;2-#

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, 321-334.

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. doi: 10.1002/(SICI)1099-0720(199810)12:5<489::AID-ACP535>3.0.CO;2-#

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,* 22, 403-435. doi:10.1016/S0273-2297(02)00005-9

Robinson, E. J. (2000). Belief and disbelief: Children's assessment of the reliability of sources of knowledge about the world. In K. P. Roberts & M. Blades (Eds.), Children's source monitoring (pp. 59-83). Mahwah, NJ: Erlbaum.

Schacter, D. L., Kagan, J. & Leichtman, M. D. (1995). True and false memories in children and adults: A cognitive neuroscience perspective. *Psychology, Public Policy, & Law,* 1, 411-428. doi: 10.1037/1076-8971.1.2.411

Templeton, L. M. & Wilcox, S. A. (2000). A tale of two representations: The misinformation effect and children's developing theory of mind. *Child Development,* 71, 402-416. doi:009-3920/2000/7102-0011

Thierry, K. L. & Spence, M. J. (2002). Source-monitoring training facilitates preschoolers' eyewitness memory performance. *Developmental Psychology,* 38, 428-437. doi: 10.1037//0012-1649.38.3.428

Thierry, K. L. & Spence, M. J. (2004). A real-life event enhances the accuracy of preschooler's recall. *Applied Cognitive Psychology,* 18, 297-309. doi:10.1002/acp.965

Table 1

Breakdown of Variables Used with Questions Used and Variable Computations

Variable	Questions used	Question Example	Variable computation
		Baseline	
Baseline accuracy	Baseline	*"What colour was the brain?"*	Number of correct baseline/Total number of baseline questions
		Training	
Recognition accuracy	Training yes/no recognition	*"Did someone tell you that their favourite frog was blue and orange?"*	Number of correct training recognition/First 4 training recognition
Source accuracy[a]	Training forced-choice source	*"Did that happened in the DVD or Real life Demonstration?"* Following a non-misleading training recognition question	Number of correct forced-choice source/First 3 training forced-choice source
		Target Interview	
Recognition accuracy	Non-misleading yes/no recognition	*"Did someone pretend a banana was a phone?"*	Number of correct non-misleading recognition/Total number of non-misleading recognition questions
	Misleading yes/no recognition	*"Did someone pretend a banana was a guitar?"* (Never occurred)	Number of correct misleading recognition/Total number of misleading recognition questions
Source accuracy	Non-misleading recognition-source combination	*"In the real-life demonstration, was there toothpaste?"*	Number of correct non-misleading recognition-source combination/Total number of non-misleading recognition-source combination questions
	Misleading recognition-source combination	*"In the DVD, was there toothpaste?"* (Happened in real life)	Number of correct misleading recognition-source combination/Total number of misleading recognition-source combination questions

Table 1 (continued)

Variable	Questions used	Question Example	Variable computation
		Target Interview	
Source accuracy	Forced-choice source	*"Did that happen in the DVD or Real life demonstration?"* (Following correct non-misleading recognition question)	Number of correct forced-choice source/Total number of forced-choice source questions
Source biases	Forced-choice source	*"Did that happen in the DVD or Real life demonstration?"* (When DVD is correct; Following correct non-misleading recognition question)	Number of correct '*DVD*' responses/Total number when '*DVD*' was correct response
		"Did that happen in the DVD or Real life demonstration?" (When real-life is correct; Following correct non-misleading recognition question)	Number of correct '*Real life*' responses/Total number when '*Real life*' was correct response
	False-alarm source	*"Did that happen in the DVD or Real life demonstration?"* (Following an incorrect '*yes*' to a recognition question)	Number of '*DVD*' responses/Total number of false alarms Number of '*Real life*' responses/Total number of false alarms

[a] Included criterion- and set-training groups only

Table 2

Means of the Proportion of Correct Recognition Training Questions

	3- to 4-year-olds		7- to 8-year-olds	
Training condition	*M*	SD	*M*	SD
Criterion	.60[a]	.30	.82[c]	.16
Set	.61[a]	.23	.84[c]	.19
Recognition	.44[b]	.24	.88[c]	.16

Note. Superscripts denote significant differences. Only the first four training recognition questions were used.

Table 3

Means of the Proportion of Correct Non-misleading Recognition-Source Combination Questions

	3- to 4-year-olds		7- to 8-year-olds	
Training condition	*M*	SD	*M*	SD
Criterion	.71[a]	.23	.82	.13
Set	.72[a]	.20	.76	.14
Recognition	.55[b]	.25	.80	.14

Note. Superscripts denote significant differences.

Table 4

Means of the Proportion of Correct Misleading Recognition-Source Combination Questions

	3- to 4-year-olds		7- to 8-year-olds	
Training condition	*M*	SD	*M*	SD
Criterion	.47	.33	.40	.33
Set	.30**	.32	.35**	.30
Recognition	.39*	.28	.36**	.32

Note. Significant values were against chance of .50.
* $p < .10$ ** $p < .05$.

Table 5

'Real life' Biases in Response to a Forced-choice Source Question after Correct Recognition

	3- to 4-year-olds		7- to 8-year-olds	
Training condition	*M*	Test statistic	*M*	Test statistic
Criterion	.63	.63	.88*	.73
Set	.61	.70	.80**	.68
Recognition	.56	.70	.77	.73

Note. Means represent the proportion of time children correctly responded *'real life'* in response to a correct forced-choice source question when the answer was *real life*. The test statistic is the mean for the proportion of the time children correctly responded *'DVD* when the answer was *DVD'*.
** $p < .001$. * $p < .05$.

Table 6

'Real life' Biases in Response to a False-alarm Forced-choice Source Question after Incorrect Recognition

Training condition	3- to 4-year-olds		7- to 8-year-olds	
	M	Test statistic	*M*	Test statistic
Criterion	.52	.48	.60*	.40
Set	.46	.54	.57*	.43
Recognition	.42	.58	.53	.50

Note. Means represent the proportion of time children responded *'real life'* in response to a false alarm forced-choice source question over the number of false alarms. The test statistic proportion of time children responded *'DVD'* in response to a false alarm forced-choice source question over the number of false alarms.

* $p < .001$.

Appendix A

Script excerpts for the Human Body event

DVD	Real-life demonstration
Now let's talk about the brain. The brain is like the boss of your body. It tells the rest of your body what to do. Without your brain you wouldn't be able to do things like kick **a soccer ball or read a book** (*kick soccer ball and show a book.)* It's always important to protect your brain when you're doing things like **riding your bike....**	Let's talk about the brain. Your brain is in your head, let's all pat our brains (*pat head)*. It feels hard because a big bone called the skull is protecting it. Your brain weighs about three pounds, that's as much as this **bag of pennies** (*Let child hold bag of pennies)* The brain looks like this (*Hold up model)*, but it feels like this **tooth paste** (*let child feel toothpaste)*
There are lots of different types of food. Some are healthy for us and good, and some are not. Here's a banana. *(Use banana as a guitar)* Look at me being silly using the **banana as a guitar**.	Now let's talk about how your body gets energy. Did you know your body gets energy from food? And you need energy to **play on swings**. When your body gets energy from food, it's called digestion.

Note. Bolded items represent episodic details children may have been asked about in the target interview. Italics represent the actions the confederate performed while reading the script.

Appendix B

Example interviewer instructions and first four training questions

Instructions: *Did learn that some people eat frogs?*

If Yes: *Did that happen in the DVD or Real-life Demonstration?* (Alternate order)

If no: *Actually someone did tell you some people eat frogs. Did it happen in the DVD or Real-life Demonstration?*

If correct source: *Good job, you remembered that you learned some people eat frogs in the DVD!*

If incorrect source: *Actually, you learned that some people eat frogs in the DVD.*

Question	**Rec Answer**		**Source Answer**	
	Child's	**Correct**	**Child's**	**Correct**
Did someone tell you their favourite frog was blue and orange?		Yes		DVD
Did you learn that some people eat frogs?		Yes		RL
Did you learn that frogs have webbed feet like ducks?		No		
Did someone put eggs into a pretend pond?		Yes		DVD

Note. Only children in the criterion- and set-training groups received source questions. Interviewers were instructed not to mention source to the recognition-training group.

Appendix C

Instructions given to children following training questions

Interviewers were instructed to say the following immediately after the training questions prior to starting the target interview:

(All children):

"You did a good job answering those frog questions. Now you know that if something didn't happen you can say no to a question.

(Only for children in the set training or criterion training condition):

And you now you know to think about whether you saw something in DVD or in the real-life demonstration.

(All children):

Now I'm going to ask you some questions about what you learned last week about the body and how it works. Remember if you didn't learn about something, or see something you can say no.

(Only for children in the set training or criterion training condition):

And try to think about where you learned things, or saw things happen."

Appendix D

Example of each question type.

Yes/no recognition question with forced-choice source

Non-misleading:

- *"Did someone kick a soccer ball?"*
- *"Was there toothpaste?"*

Following a correct *'yes'* response: *"Did that happen in the real-life demonstration or DVD?"*
If incorrect *'no'* move on to next question.

Misleading (Items that never happened):

- *"Did someone throw a tennis ball?"*
- *"Was there a wrinkly sponge?"*

Following a correct *'no'* response move onto the next question.
If incorrect *'yes' "Did that happen in the DVD or real-life demonstration?"*

Recognition-source combination

Non-misleading:

- *"In the DVD, did someone kick a soccer ball?"*
- *"In the real-life demonstration, was there toothpaste?*

Misleading:

- *"In the real-life demonstration, did someone kick a soccer ball?"* (Happened in DVD)
- *"In the DVD, was there a wrinkly sponge?"* (Happened in the real-life demonstration).

Note. Example items taken from script provided in Appendix A. Children were asked about each of the 32 questions in one of these formats. Items were counterbalanced across conditions to be asked in both non-misleading and misleading ways throughout.