Wilfrid Laurier University

Scholars Commons @ Laurier

Theses and Dissertations (Comprehensive)

2017

Young Children's Source Monitoring: Exploring the Contexts of Task Difficulty and Repeated Events

Becky Earhart
earh9100@mylaurier.ca

Follow this and additional works at: https://scholars.wlu.ca/etd

Part of the Developmental Psychology Commons

Recommended Citation

Earhart, Becky, "Young Children's Source Monitoring: Exploring the Contexts of Task Difficulty and Repeated Events" (2017). *Theses and Dissertations (Comprehensive)*. 1926.
https://scholars.wlu.ca/etd/1926

This Dissertation is brought to you for free and open access by Scholars Commons @ Laurier. It has been accepted for inclusion in Theses and Dissertations (Comprehensive) by an authorized administrator of Scholars Commons @ Laurier. For more information, please contact scholarscommons@wlu.ca.

Young Children’s Source Monitoring: Exploring the Contexts of Task Difficulty and Repeated Events

by

Becky Earhart

Master of Arts, Wilfrid Laurier University, 2012

Dissertation

Submitted to the Department of Psychology

in partial fulfillment of the requirements for

Doctor of Philosophy in Developmental Psychology

Wilfrid Laurier University

© Becky Earhart, 2016

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

This dissertation had two over-arching goals. The first was to study the cognitive mechanisms underlying effective source monitoring by clarifying the role that developing executive function skills play in children's increasing ability to monitor sources. The second goal was to examine whether a particular interview technique called "source-monitoring training" could help children to recall the sources of their memories more accurately. These two separate lines of research were furthered by the same methodology, and thus, these separate research questions were examined simultaneously within both of the experiments conducted for this dissertation.

In the first study, the difficulty of the source-monitoring decisions was manipulated by testing 4- to 8-year-old children's memories of a lab-based event after a shorter delay (1-2 days) or a longer delay (8-10 days). Within these two conditions, I explored both the relationship of source monitoring to executive function, as well as the effectiveness of the source-monitoring training procedure. The results of this study showed that executive function was related to source monitoring, and mediation models demonstrated how children's source monitoring improves with age due to developments in working memory, which improve event encoding and therefore, source monitoring. The effects of source-monitoring training were not as clear as expected; the only group to benefit from the training were older children in the shorter delay condition. Interestingly, neither the relationship between executive function and source monitoring nor the effects of source-monitoring training were affected by the difficulty of the task in the way expected.

In the second study, 4- to 8-year-old children's source monitoring was examined within a repeated-event paradigm. The inclusion of more than two sources (i.e., events) created a more

realistic and generalizable task. Again, both the relationship between executive function and source monitoring and the effectiveness of source-monitoring training were examined within the same study. In this study, there was evidence that two broad components of executive function as measured through parent reports were related to source monitoring. The source-monitoring training did not improve source accuracy, but did impact the types of errors children made, such that older children who received the training were more likely to say, “don’t know” instead of confusing the events.

Testing these relationships in a variety of conditions illustrates how cognitive and interview factors are related to source monitoring, demonstrating clear links between executive function and source monitoring, but mixed evidence for the effectiveness of source-monitoring training. Collectively, my doctoral program of research contributes a greater understanding of how source monitoring develops and whether source-monitoring training could be used in practice.

Acknowledgements

This acknowledgments section could probably be longer than the dissertation itself. There are way too many thank-yous to include here because my tribe is amazing, but I'll give it my best shot.

This degree and this process more generally have been a self-discovery journey for me, in which I have learned a lot about what I am capable of, and what I want to spend my life doing. I have spent the last six years at Laurier working extremely hard to make an impact on the research community and the world outside of research.

I owe a great deal of my success to my mentors over the years, who taught me the value of hard work, critical thinking, and collaboration. From my official supervisor, Kim Roberts, to my unofficial supervisors and colleagues, Sonja Brubacher, David La Rooy, Michael Lamb, and Laura Melnyk, each has had a unique impact on my learning and my accomplishments. Most importantly, they have instilled in me a sense of responsibility to build on the legacy I have inherited. Coming after a line of such brilliant researchers is an honour, and I will always be sure to open doors for others the way that they have for me. In particular, a special thank you to Sonja, whose support through the difficult times during my PhD has meant more to me than she will ever know. I can truly say I wouldn't be where I am today without her.

I would also like to express my heartfelt gratitude to my committee members, Alexandra Gottardo and Eileen Wood. Besides their significant contributions and feedback on my dissertation, they both provided emotional support and guidance that was invaluable to me as I worked through my dissertation and figured out what would come after the PhD. It was an absolute pleasure working with both of them.

I am so grateful to my friends and family, the amazing people in my life who have lived this with me and experienced everything that I have put into it. It can't have been easy, but your support has meant everything to me. To two special labmates, Courtney Arseneau and Leanne Bird, who have become forever-friends: I am so thankful that you were by my side during these years. You are both inspirational young women that challenge me, understand me, and bring a little sunshine into my life when it’s needed most. One of the best parts of this journey was meeting you along the way.

To my partner, Wyatt. I want to thank you for bringing balance into my life; for your constant support and unconditional love; for reminding me not to take life too seriously; for your patience with my long work hours and other workaholic tendencies; for being a safe place to land at the end of a hard day. For everything you do, I can’t thank you enough. Looking forward to many more milestones with you.

I would like to thank the many research assistants that volunteered their time to work on various projects with me. It was such a pleasure working with each and every one of you and I hope we’ll meet again. Thanks also to the hundreds of children and their parents who participated in the research conducted for this dissertation. It would not have been possible without your help.

This degree is not an end point for me, but a beginning. It is a symbol of the university’s trust in me to go out into the world and continue proving the worth of my degree and my time at Laurier. Although the PhD is a huge accomplishment, I still have so much to learn! And as I continue to learn and grow, I’m happy to have such wonderful people who have supported me through thick and thin by my side.

Young Children's Source Monitoring: Exploring the Contexts of Task Difficulty and Repeated Events

General Introduction

Memories of events come from a variety of sources; for example, real-life experiences versus events seen on television. Source monitoring is the process of making decisions about the sources of memories (Johnson, Hashtroudi & Lindsay, 1993). For example, reasoning about whether information came from speaker A or speaker B, or was directly experienced versus suggested by another person.

Source monitoring has many social, educational and forensic applications. From a social perspective, effective source monitoring may help one avoid embarrassment, like the awkwardness that arises when you tell a joke to the same person who told it to you in the first place. Another source-monitoring problem that occurs in social contexts is incorporating fictional narratives into one's knowledge base as facts (Johnson et al., 1993). Stories that people hear from others can be confused with things that they have directly experienced themselves or learned from a more factual source, and this could lead to the spread of erroneous information through social interactions between uninformed sources.

From an educational perspective, source monitoring is important because encouraging children to think critically about sources of information they encounter in everyday life could help them distinguish between credible and "incredible" sources (i.e., a source that is non-credible; e.g., a teacher versus an unreliable website; Robinson, 2000). Learning occurs in a variety of contexts, and as technology becomes commonplace in the classroom, children have access to a vast number of sources, some of which may not be reliable. It is important to teach children how to evaluate the credibility of different sources, which depends on first identifying the source of information. Source-monitoring research also addresses questions such as how children learn from different sources, how they integrate information from multiple sources during learning, and how they edit their knowledge base when they discover that a source they learned from is not credible (e.g., Renner & Roberts, 2010).

Finally, source monitoring is relevant to forensic investigations because a witness may be asked many questions about the context (or source) of an event. For example, where you were, who you were with, and when an event occurred are all aspects of the source of a memory, and are important details in investigations. Challenges associated with recalling accurate source details are especially relevant when it comes to children’s testimony because of the difficulty that children have with source monitoring (Roberts, 2002). In some cases, such as allegations of child abuse, children’s testimony may be the only evidence available and thus, it is important to obtain the most complete and accurate information possible.

In criminal investigations of abuse, children are often asked to recall events that have happened on multiple occasions because child abuse is often repeated (Trocmé et al., 2010). In legal systems derived from British Common Law (e.g., in Canada and Australia), a child is required to describe details specific to one incident so that specific charges can be laid, and so that the defendant has an opportunity to challenge the charges (Guadagno, Powell & Wright, 2006). Describing one occurrence of abuse requires monitoring the source of memories in order to avoid confusing it with other incidents (i.e., reporting details from a different occurrence of abuse). Children may confuse details from similar events or even incorporate things they have seen on television into their reports because they confuse the origins of their memories (Connolly & Lindsay, 2001; Roberts & Blades, 1999). Therefore, a complete understanding of children’s source-monitoring abilities is essential to giving children the best chance at providing accurate testimony, which may have a positive impact on the currently very low prosecution rates in these cases.

Roadmap

The following research examined aspects of cognitive skill and the interview situation that may promote increases in children’s accuracy for distinguishing between multiple sources. The research focused on three issues. First, each study examined cognitive abilities related to executive functioning that may contribute to source-monitoring development across childhood. Second, the studies examined an interview technique that may help children overcome their difficulties with source monitoring. Third, the research focussed on differences from early to mid-childhood to assess developmental changes. The following sections set the context for the research questions by describing the origins of this field of research, providing a theoretical overview of source monitoring, describing source-monitoring development and putting it in the greater context of cognitive development from age 3 to age 8, and finally, discussing current research on interventions aimed at improving children’s source-monitoring accuracy.

Origins of this Field of Research

In the 1980s and early 1990s there were a number of highly publicized cases of alleged mass child abuse occurring in daycares. The accusations in these cases often involved satanic and ritualistic child abuse of many children at the daycares, and the details of the cases were both horrifying and bizarre. These infamous cases raised concerns about the interviewing methods used to collect evidence from the children and the impact that these techniques may have had on the quality of the children’s testimony (Lamb, Orbach, Hershkowitz, Esplin & Horrowitz, 2007; Bruck & Ceci, 1999). One of these cases is notable because of the involvement of psychologists who provided information to the court about interviewing child witnesses.

Specifically, in April of 1985, a daycare worker named Margaret Kelly Michaels was accused of abusing children at Wee Care Nursery School in New Jersey, USA, where she worked. The accusations began with a comment from a child who had his temperature taken using a rectal thermometer during a visit with the doctor and stated that Kelly had done the same thing at daycare. Eventually this comment was reported to authorities, and what followed was a series of poorly conducted interviews with the 50 children that attended the daycare. These interviews involved techniques such as using peer pressure, asking children to speculate about what might have happened, and bribing the children in exchange for statements against Michaels. The children that were interviewed alleged such things as being penetrated with forks and knives and being forced to eat human excrement. The case went to trial in June of 1987 and despite the lack of physical evidence to corroborate the children's claims, Michaels was convicted of 115 counts of sexual offences and was sentenced to 47 years in prison (Rosenthal, 1995).

Michaels appealed and the decision was reversed after she had served 5 years of her sentence. The State attempted to re-try Michaels, but was prevented from doing so because the reliability of the children's testimony had been called into question. Two developmental psychologists, Maggie Bruck and Stephen Ceci, wrote an amicus brief to the court to discuss some of the issues in the case and describe relevant research on children's memories. The brief highlighted the role of interviewer bias, repeated questions, peer pressure, and the use of anatomically detailed dolls in contaminating the children's reports, and concluded that these techniques could have led to memory errors or false memories (Bruck & Ceci, 1995). This case was influential in initiating further investigation regarding the reliability of children's testimony and the treatment of child witnesses in the legal system.

In the aftermath of the daycare cases it became clear that further research was necessary to discover the strengths and limitations of young children's eyewitness testimony. This prompted an exponential increase in research on children's suggestibility in the early 1990s (Lamb et al., 2007; Bruck & Ceci, 1995); that is, the idea that children's memories and reports can be shaped by suggestions made to them by interviewers (Bruck & Ceci, 1999). Suggestibility has been thoroughly examined with different age groups, and researchers have also studied individual difference factors (see Bruck & Ceci, 1999; Bruck, Ceci & Melnyk, 1997; and Melnyk & Bruck, 2004, for reviews). Based on this large body of research it became clear that interview techniques can strongly affect how children respond.

Researchers then turned toward the goal of developing evidence-based recommendations for front-line workers that could help children report their memories more accurately. In the last 25 years researchers have generated a strong consensus about the basics of interviewing and child development. For example, based on a large literature of research findings, it is widely recommended that interviewers use open-ended questioning (e.g., "Tell me what happened") as much as possible to improve the quality of children's reports (e.g., Hutcheson, Baxter, Telfer & Warden, 1995; Oates & Shrimpton, 1991; Ornstein, Gordon & Larus, 1992). In addition, building rapport and outlining "ground rules" at the beginning of an interview are included in interviewing protocols to help interviewers maximize the accuracy and completeness of children's reports (Lamb et al., 2007). The development of well-researched protocols with explicit interview strategies has led to a more positive view of how children can participate in the legal system. In particular, the focus has shifted from examining conditions which make them unreliable, to what they are capable of contributing (Pipe, Lamb, Orbach & Esplin, 2004). Children are often able to provide valuable information in interviews, but it is essential that these interviews involve careful investigative procedures that align with children's developmental capabilities.

With a basic understanding of the most effective interview practices well developed, researchers in the field are now turning back to basic questions about how memory works and develops, including what cognitive factors underlie developmental gains in memory processes. As well, researchers continue to explore the impact of various interview-related variables on children's reports to expand our knowledge of techniques that may improve the quality of children's testimony. This dissertation contributed to both of these broader goals in the current literature by answering basic questions about developmental memory processes (i.e., the cognitive factors underlying source-monitoring development) as well as examining the impact of a specific interview technique (i.e., source-monitoring training) on the accuracy of children's reports.

Defining Source Monitoring

A source refers to the conditions under which a memory was acquired (Johnson et al., 1993). This could be aspects such as the time, place, or media through which it was experienced (e.g., a real experience, a dream, something you imagined, or something you saw on television). Source monitoring is the process of making decisions about the origin of known or remembered information (Johnson et al., 1993). In everyday life people constantly monitor the sources of episodic memories (e.g., have I actually gone to Niagara Falls or did I see it on TV?), as well as factual information (e.g., Did I read it in a peer-reviewed paper or on Wikipedia?). Source monitoring is important for many cognitive functions and affects our everyday beliefs, opinions and behaviours. Differentiating sources is essentially the key to the phenomenological experience of remembering because if memories are retrieved without contextual information, they are experienced as semantic knowledge rather than episodic memories (Johnson et al., 1993; Tulving, 1984).

Researchers distinguish between several types of source-monitoring judgments depending on whether the sources involved are external or internal to individuals (i.e., information derived through perceptual processing of external sensory properties, versus internal self-generated thought processes such as imagination). An external source-monitoring task is one of deciding between two external sources; for example, whether information came from Speaker A or Speaker B. An internal source-monitoring task involves distinguishing between two internal sources; for example, whether you said something aloud or just thought it to yourself. Reality monitoring is a term used to describe source judgments involving both external and internal sources; for example, did someone tell you a story or did you dream it? (Johnson & Raye, 1981; Johnson et al., 1993). Adults are more accurate with reality monitoring than with external source monitoring, which shows that internal and external sources function in different ways; the cognitive operations (such as organization and elaboration) associated with generating information serve as cues to the source of that information at retrieval (Raye & Johnson, 1980).

Two factors that are known to impact the ease and accuracy of source-monitoring decisions are the delay between an event and recall and the similarity of sources. The negative effect of delay on episodic memory is well-documented, and the same is true regarding memory for source. For example, several studies examining misinformation effects at 10-minute, 1-week, and 1-month intervals show that adults are more likely to accept misinformation at longer delay intervals because of greater confusion about the original source of the information (Underwood & Pezdek, 1998; Frost, 2000; Frost, Ingraham & Wilson, 2002). The longer the delay, the more difficult it is to make source decisions, and accuracy decreases in kind. There is also a substantial literature examining the effect of source similarity, demonstrating that the more similar sources are the more difficult it is to distinguish between them (e.g., Lindsay, Allen, Chan & Dahl, 2004; Day, Howie & Markham, 1998; Roberts & Blades, 1999; Foley, Harris & Hermann, 1994). Sources that are more similar have fewer distinct or unique cues that can be used at retrieval to attribute source information (Roberts, 2002).

It is clear that source monitoring is difficult and that source errors can (and frequently do) occur. Relevant theories of source monitoring can help explain how source decisions are made, and why factors such as delay and source similarity increase the likelihood of source errors.

Theoretical Models of Source Monitoring

Several theoretical models are relevant in explaining the cognitive processes underlying source monitoring. The two main theories that are discussed in the context of this dissertation are Fuzzy-Trace Theory and the Source-Monitoring Framework. Fuzzy-Trace Theory lends itself more to explaining how source information is encoded and stored in memory, whereas the Source-Monitoring Framework is more explanatory in terms of how source decisions are made.

Fuzzy-Trace Theory. Fuzzy-Trace Theory explains how memories are formed and how they decay over time. The theory postulates that dual representations of experiences are encoded in parallel; gist and verbatim. Gist traces are vague representations of the general sense or pattern of what is being encoded, including the meaning or structure of an event. Verbatim traces, on the other hand, represent the content of memories by preserving surface details exactly (verbatim). Most of our remembering occurs in gist form because gist representations are more accessible in memory and require less effort to retrieve (Brainerd & Reyna, 1990). In addition, verbatim traces decay more rapidly than gist traces, so it becomes more likely with the passing of time that gist traces will be retrieved because verbatim information about particular experiences may be lost (Brainerd & Reyna, 1995). Memories become more generalized and less detailed over time.

It is argued that source is encoded and represented as a verbatim trace. If the verbatim trace is still well integrated in memory and accessible, source decisions are made through direct retrieval when the information is cued (Brainerd & Reyna, 2004). Since verbatim traces decay faster than gist traces, source information is lost early on while the gist of an experience is retained longer. This helps to explain why delay negatively affects source monitoring. After long delays it is more likely that source information is lost and cannot be directly retrieved even though the event itself can be remembered. Source confusions can occur when the verbatim traces containing source information have decayed, and we instead accept information that is consistent with the gist, although it may not be correct (Thierry, Spence & Memon, 2001). Gist interference with verbatim traces is more common after a delay (Titcomb & Reyna, 1995).

The Source-Monitoring Framework. The Source-Monitoring Framework, proposed by Johnson and colleagues more than 20 years ago, is a theory that seeks to illuminate the cognitive process of source monitoring by explaining how judgments about source are made and what criteria are used for those judgments (Johnson et al., 1993). Fundamental to the Source-Monitoring Framework is the idea that source monitoring can involve making attributions about the origin of memories, which is more complicated than simply retrieving a memory trace that specifies source information. Source monitoring also involves the use of complex decision-making processes based on retrieved information (Johnson et al., 1993).

According to the Source-Monitoring Framework, there are two ways of making source decisions. The first is through *heuristic judgment processes*, which involve quick decisions that may occur in the course of remembering without conscious awareness of the decision-making process (e.g., you immediately recall the person's voice and conclude that that person was the source; Johnson et al., 1993). These decisions are based on the qualitative characteristics of memory traces, such as the spatial or temporal context, the amount of perceptual detail, the cognitive operations associated with the memory, semantic details, and the affective response from when the memory was formed. Decisions can be made by comparing differences in the characteristics of memories from different sources. For example, when distinguishing between an event that actually happened and something that was imagined, a real event would have more perceptual detail associated with the memory, whereas an imagined event would have minimal or no perceptual detail, and instead contain details about cognitive operations, such as organization and elaboration (Johnson et al., 1993). This theory provides a strong explanation for source similarity effects because when sources are highly similar, there is more overlap between the characteristics of the sources. Therefore, distinguishing between them is more difficult because there are few unique cues to identify sources.

Some source decisions require a more deliberate, analytic approach through what is called *systematic judgment processes*. When making decisions this way people reason carefully about possible sources, which may involve retrieving supporting memories, reasoning about constraints, and employing strategies (Johnson et al., 1993). For example, when trying to remember who told you a joke, you might recall that you were at work when you heard the joke so it must have been a co-worker who told it. This decision-making process requires retrieving supporting memories about where you were when you heard the joke in order to reason about possible sources. Johnson and colleagues provide indirect support for systematic processes by citing research that connects deficits in source monitoring with frontal lobe damage or dysfunction – the same brain regions that are implicated in higher order cognitive functions such as reasoning (Johnson et al., 1993; Schacter, Kagan & Leichtman, 1995). This process is clearly more complicated than simple retrieval of source information, as proposed by Fuzzy-Trace Theory.

Making source decisions involves setting a criterion for making a judgment and comparing the retrieved information to that criterion. This could involve determining which characteristics are most important for the decision and how confident one feels about those characteristics. If the confidence level surpasses the criterion the memory will be attributed to that source. Criteria can be made more or less stringent depending on a number of situational factors, such as goals or motivation (Johnson et al., 1993). Empirical evidence for this concept comes from studies demonstrating that people are less suggestible if provided with incentives for correct responses or penalties for incorrect responses (e.g., Roebers & Schneider, 2005, Roebers, Moga, & Schneider, 2001, Koriat & Goldsmith, 1996). For particularly important source decisions, such as giving evidence in a forensic investigation, the criteria may be more stringent to increase the likelihood that sources are attributed correctly.

Comparing Theories of Source Monitoring. It is clear that source-monitoring decisions depend on the quality of both the episodic memory and the decision-making process. Fuzzy-Trace Theory focuses on the quality of the memory traces and provides an explanation of the structural representation of source information in memory, whereas the Source-Monitoring Framework highlights the important role of how the decision-making process occurs. These theories differ in terms of how they propose source information is accessed from memory; according to Fuzzy-Trace Theory source information can simply be retrieved, whereas the Source-Monitoring Framework proposes a dynamic decision-making process involving reasoning and strategy.

If source information has been encoded and retained (i.e., the verbatim trace has not decayed), the complex decision-making processes described in the Source-Monitoring Framework may not be necessary. And, in fact, heuristic processing by comparing characteristics as described in the Source-Monitoring Framework may be experienced as simple "retrieval" if one is not applying considerable effort in making that decision. In that sense, there may be some alignment between the theories in describing how simple, effortless source decisions occur. However, people often do not pay attention to the sources of their knowledge, which means that the information is not accessible through a simple retrieval process (e.g., Jacoby, Kelley, Brown & Jasechko, 1989; Marsh, Landau & Hicks, 1997). Therefore, it is likely that people will be required to engage in reasoning processes to make source decisions, and the Source-Monitoring Framework adds a decision-making component that allows for the reconstructive nature of memory. Fuzzy-Trace Theory does not include a mechanism to describe how source decisions are made in this context.

The Development of Source Monitoring

Extensive developmental research has shown that source monitoring is difficult for children but develops gradually across childhood, with the largest improvements between age 3 to age 8 (e.g., Poole & Lindsay, 2001; Roberts & Blades, 1999; see Roberts, 2002, for a review). Children typically acquire an implicit understanding of sources before they can explicitly report source information. For example, even children as young as 3-years-old trust and report information from informative sources more than uninformative sources (e.g., Scofield & Behrend, 2008; Robinson, Butterfull & Nurmsoo, 2011;), indicating that they can differentiate between sources. However, they often cannot report the source of their beliefs (Whitcombe & Robinson, 2000) or explain how they know whether or not a source is reliable (Roberts, 2002).

Although there are some situations where even young children perform as well as adults (e.g., discriminating between something you have said versus something you have only thought; Foley, Johnson & Raye, 1983), it is not until approximately age 10 that children perform as well as adults on many source-monitoring tasks (Roberts, 2002; Roberts & Blades, 1996).

According to Fuzzy Trace Theory, source errors stem from the loss of verbatim traces that contain source information. Relevant to children's source-monitoring development, then, is the fact that children lose verbatim traces faster than adults do (Brainerd & Reyna, 2004). This means that source information is lost more quickly for children. The Source-Monitoring Framework, on the other hand, highlights aspects of systematic processing such as reasoning and heuristic strategies. Children may not have the cognitive skills required for complex reasoning such as retrieving supporting memories and holding them in mind while making a decision. Children may also struggle with selecting an effective strategy for the task, or fail to benefit from the strategy that they select.

As was noted earlier, there are several factors that can make source monitoring more difficult for adults (i.e., delay between event and recall, and source similarity). When there are challenging conditions for source monitoring this has an even greater detrimental effect on children's source accuracy than it does for adults. Ackil and Zaragoza (1995) found that 6-, 8- and 10-year-olds made significantly more source errors by accepting misinformation after one week compared to an immediate source-monitoring test. Delay affects children's reports even more than adults' because, as discussed above, they lose verbatim traces faster than adults do (Brainerd & Reyna, 2004).

In terms of similarity, Roberts and Blades (1999) had 4- and 10-year-old children watch a live event and a video that was either similar or different from the live presentation. One week later, children in the similar condition were more confused than those in the different condition, and made more source-monitoring errors by reporting details from the video as having happened in the live event. Research has also shown that the strength of the source similarity effect depends on the age of the participants. Lindsay, Johnson and Kwon (1991) manipulated source similarity by using voices of the same gender or different genders presented to the left and right side of the participant. They found that 4-year-olds had far more difficulty than adults when the voices were of the same gender (more similar) than when they were different genders (less similar) - the similarity effect was exaggerated for young children compared to the adult group.

Children under 10-years-old may have more difficulty discriminating between two similar sources than adults do because adults have a greater ability to think about different dimensions of source. If two sources are highly similar on one dimension (e.g., the gender of the speaker), one may need to consider another dimension in order to distinguish between them (e.g., left or right presentation). Whereas adults may have the ability to think about more than one dimension concurrently while they work on this problem, children may not be able to do so (Brocki & Bohlin, 2004). Another explanation provided by Lindsay and colleagues (1991) was that when the sources are distinct and judgments are easy, young children perform comparably to older children and adults because very little strategizing is required; however, when the task is difficult, a strategy is required and children may not have the cognitive skills or metamemory to produce an effective strategy (Lindsay et al., 1991).

The problems that children have with delay and source similarity demonstrate the role of cognitive issues in source-monitoring development. It is clear that children struggle more than adults do with source monitoring. Identifying the age-related factors that contribute to these difficulties is best examined within the context of cognitive development more broadly between the ages of 3 and 8.

Cognitive Skills Underlying Effective Source Monitoring

One of the earliest developing cognitive skills necessary for accurate source monitoring is improvements in episodic memory. In order for the characteristics of a memory to be examined (and for a subsequent source attribution to be made), the event must first be remembered. Although children can remember events in their lives after short delays by age 2 (Peterson & Rideout, 1998), they may not begin to monitor the sources of their memories until much later.

At age 3, children may not be able to justify why they know something (O'Neill, Astington & Flavell, 1992; Waters & Beck, 2012). They are preoccupied with expanding their knowledge, and do not pay careful attention to the sources of knowledge. Children tend to view all adults as highly credible sources (Jaswal, Carrington Croft, Setia & Cole, 2010), and therefore, it is not important to remember from which adult they learned information. Young children would not have enough experience making source decisions to understand why it is important to discriminate between sources, and hence, may not pay particular attention to source at encoding.

Eventually children begin to understand that knowledge is connected to different sources, and that one must have access to certain informational sources in order to gain knowledge. For example, in order to know what colour an object is, one would need to see it; to judge how heavy it is, one would need to feel it. By age 4 to 5, children can explain how they know what is in a container (e.g., because they have seen it or because they were told; Tang & Bartsch, 2012). As children come to understand that there are different sources of information, the foundation of source monitoring is available.

Around the same time that children learn that knowledge is connected to different sources, they also begin to develop a theory of mind; that is, an understanding of others' mental states and how actions are influenced by mental states (Wellman, Fuxi & Peterson, 2011). Theory of mind has been related to source monitoring or suggestibility in several studies (e.g., Bright-Paul, Jarrold & Wright, 2008; Welch-Ross, 2000; Welch-Ross, Diecidue & Miller, 1997). Understanding that people can have different representations or beliefs about the same events helps children to avoid accepting misinformation. They also become more aware that because there are different sources of information that may hold different beliefs, some sources may be more credible than others. However, this does not mean that children can accurately monitor source. As discussed above, children's cognitive limitations may prevent them from carrying out effective source-monitoring processes. One key factor involved in children's ability to engage in higher order cognitive processes is executive function.

Executive Function. Executive function is a broad category of skills that support planning and goal-directed behaviour (DeLuca & Leventer, 2008; Zelazo, Muller, Frye & Marcovitch, 2003; Zelazo & Muller, 2002). There is still debate in the literature about the structure and components of executive function (e.g., whether there are two or three, or possibly more, factors), but two components that are widely agreed upon are inhibitory control and working memory. Inhibitory control is the ability to ignore information that is not relevant to the current task and restrain automatic responses (Roberts & Powell, 2005b). Working memory allows for temporary storage and manipulation of information in order to complete complex cognitive tasks (Baddeley, 1992). Both inhibitory control and working memory develop throughout childhood, and there are concurrent improvements in source monitoring. There are theoretical reasons to believe that both inhibitory control and working memory would be necessary for source monitoring, and may therefore contribute to children's source-monitoring development.

Inhibitory control would be required to inhibit familiarity-based retrieval processes that are often used automatically to make recognition decisions (Ruffman, Rustin, Garnham & Parkin, 2001). Higher levels of inhibitory control would also allow children to ignore information from competing sources in order to make a correct source judgment; for example, reporting information about one instance of a repeated event while inhibiting reporting details from other similar events. Working memory would be highly involved in systematic judgment processes because this type of decision requires strategy use and the retrieval of supporting memories. In order to do this, children would be required to hold this extra information in mind while making a decision. Working memory also plays a role in controlling attention, and therefore, designates to what information cognitive resources will be allotted (Gerrie & Garry, 2007). A complex process of reasoning about the constraints of memories, retrieving supporting memories, comparing and contrasting sources, and inhibiting competing information may be needed to make effective decisions about source.

The current literature on executive function and source monitoring in children is not extensive. Research generally tends to show that executive function is related to both episodic memory and source monitoring. However, the results are rarely that simple, often involving qualifications about complex relationships. In a comprehensive review of individual differences in suggestibility, Bruck and Melnyk (2004) found that only half of studies showed significant correlations between executive function and suggestibility (a particular type of source-monitoring error); those studies that did find significant relationships demonstrated that increased executive function was positively related to resisting misinformation. Similarly, Roberts and Powell (2005b) found that children with better inhibitory control were more likely to resist suggestions, and Karpinksi and Scullin (2009) replicated those results with preschoolers, as well as showing a relationship with working memory. However, several researchers have found mixed results, such as showing one component of executive function to be related to source monitoring but not another component, or showing a relationship with one type of source-monitoring task but not another (Melinder, Endestad & Magnussen, 2006; Ruffman et al., 2001).

Overall, the results of this literature provide support for a relationship between executive function and source monitoring, but also show that the relationship is complex and seems to vary with the task demands, highlighting the need for further research that can explain the differences in outcomes across studies. It is likely that other factors relating to differences in methodology are influencing the strength of these relationships in various studies. Therefore, this dissertation examined the relationship of executive function to source monitoring in a variety of conditions including easier and more difficult tasks, and tasks involving two external sources as well as tasks involving many sources (i.e., a series of repeated events). By isolating individual factors such as task difficulty that may affect whether executive function and source monitoring are related, the present research addressed potential methodological issues that may account for the mixed results of previous studies in this area.

Using Strategies for Source-Monitoring Decisions. Once the cognitive structures required for effective source monitoring are in place, children need to develop strategies that are helpful for source monitoring so that they can use those newly-developed cognitive skills in a successful way. Examples of strategies that could be used to aid in source monitoring include retrieving supporting memories, reasoning about the constraints surrounding possible sources, comparing and contrasting the characteristics of different sources, or setting criteria that are more or less stringent depending on the importance of the source decision.

Research on children's strategy use shows that young children may fail to produce an effective strategy for decision-making (a production deficiency), or they may use a strategy that does not benefit their performance (a utilization deficiency; see Bjorklund, Miller, Coyle & Slawinski, 1997, for a review). A strategy may not be effective due to a lack of background knowledge, a lack of resources available in working memory, or even a lack of motivation to carry out the strategy effectively. Whether children fail to produce an appropriate strategy or fail to benefit from it, , age is an important consideration. Generally younger children are less effective at using strategies compared to older children or adults (Bjorklund et al., 1997).

With respect to source monitoring, young children may not have had enough practice making source-monitoring decisions to be aware of the qualitative characteristics that they can use to compare different sources. For example, children might not be aware that memories high in perceptual detail are more likely to have been experienced directly, whereas vague memories that lack perceptual detail were probably experienced through another media. Children's failure to select an appropriate strategy, such as comparing sources based on perceptual detail, is explained by a lack of metamemory (Roberts, 2002). In particular, children have little awareness of how their memory works or what strategies they could use. This makes source monitoring more difficult for children because they do not narrow their focus to useful differences between sources that can help to distinguish between them. In cases where children do have the cognitive skills necessary for source monitoring but demonstrate a production deficiency with regard to a strategy, instructions in strategy use or direct facilitation of a strategy may improve source accuracy. Several studies that attempt to improve children's source-monitoring skills through interventions targeting strategy use are discussed below.

Interventions Aimed at Improving Source Monitoring

Source monitoring has many applications, and particularly because of the significance of these applications in forensic settings, it is important to discover ways to improve children's source-monitoring accuracy. Recent research has focussed on factors surrounding the way interviews are conducted to determine what, if any, interview techniques could help improve source accuracy in children's reports.

Earhart and Roberts (2014) examined the impact of facilitating different recall strategies during a memory interview to improve source-monitoring performance. This work was theoretically grounded in the Source-Monitoring Framework, which, as discussed above, postulates that decisions are made by comparing the characteristics of memories to determine which source fits best with a memory (Johnson et al., 1993). It was predicted that asking children to consider information from two sources at the same time would facilitate a strategy of comparing sources, and, therefore, lead to more accurate source-monitoring scores than would asking children to consider sources one at a time in a serial fashion.

To test this prediction, Earhart and Roberts (2014) had interviewers ask children to recall information from two different sources either serially (i.e., information from one source at a time) or in parallel (i.e., information from two sources simultaneously). Accuracy did not differ between these two conditions for the older children (7- to 8-year-olds) who were likely proficient in producing and implementing effective strategies in both conditions. However, for the younger children (4- to 6-year-olds), who likely needed assistance with strategy use in relation to source monitoring, there were significant benefits in the parallel condition. These younger children, who may not have been cognitively ready to produce or implement strategies of their own, benefitted from the facilitation of a compare and contrast strategy. This is one example of an interview technique that promotes accurate source monitoring with even the youngest age group of children by enhancing strategy use.

Source-Monitoring Training. Another interview technique called "source-monitoring training" also targets strategy use as a means to improve source-monitoring accuracy. The training procedure involves providing participants with practice in a source-monitoring task prior to conducting a memory interview. The typical paradigm involves a laboratory event with exposure to two or more sources. After a delay children receive training through a practice source-monitoring task with unrelated stimuli, and immediately after training children complete a memory test about the sources from the event. Several recent studies using this paradigm have found that children can be trained to monitor sources more accurately (e.g., Poole & Lindsay, 2002; Thierry & Spence, 2002). Notably, in many of these studies the children are asked about sources in parallel during the source monitoring test, which means that the training technique demonstrates benefits above and beyond structuring an interview to facilitate parallel processing, as found by Earhart and Roberts (2014).

Researchers suggest that the training works by drawing attention to source information as task-relevant and encouraging or improving strategy use (e.g., Poole & Lindsay, 2002; Thierry & Spence, 2002; Thierry et al., 2001). In part, the interview technique increases accuracy because people think more carefully about sources and use stricter criteria for their decisions when they know that source information is important (Thierry & Spence, 2002; Koriat & Goldsmith, 1996). This technique shows promise, but there are inconsistencies in the literature about which age groups benefit from training and there are still unanswered questions regarding situational factors that may influence the effectiveness of the training procedure.

As research on source-monitoring training developed, the methodology that was used changed considerably. Many early studies of source-monitoring training did not include a non-training control group for comparison (e.g., Poole & Lindsay, 2001; Giles, Gopnik & Heyman, 2002). Of those that did, some used a more implicit form of source-monitoring training where children were simply asked source-monitoring questions about target events before providing a free recall account (Thierry et al., 2001; Leichtman, Morse, Dixon & Spiegel, 2000). Others included explicit feedback about sources in the training procedure and used non-target sources for the training task in order to measure the transference of the training effect (Poole & Lindsay, 2002; Thierry & Spence, 2002). Some studies only used a single age group, so these studies are less informative in terms of developmental differences in the effects of source-monitoring training (Giles et al., 2002; Thierry, Lamb, Pipe & Spence, 2010). The following discussion will focus primarily on those studies that included a control group and examined effects amongst children of more than one age group.

Poole and Lindsay (2002) studied source-monitoring training by having 3- to 8-year-old children interact with “Mr. Science” (a research assistant who conducted science activities with the children), and then hear misleading stories about the activities 3 months later. Training was provided prior to the target interview; a research assistant acted out some actions and talked about others, and children were asked about which actions were actually done and which were only mentioned. Children were given feedback on their responses. The 7- and 8-year-olds were less likely to provide false information about the Mr. Science activities in the interview, but for the 3- to 6-year-olds there was no benefit of training. One reason why the younger children may not have shown a training effect in this study is that the delay was three months; younger children may be more susceptible to forgetting over time, so perhaps they had weaker memory traces for the event and not even training could help them monitor source more effectively (Poole & Lindsay, 2002). Training may not work for young children when the task is extremely difficult, and such a long delay would make this task very difficult for the younger age group.

Thierry and colleagues (Thierry et al., 2010; Thierry & Spence, 2002; Thierry et al., 2001) conducted several studies using a similar science activities paradigm to Poole and Lindsay's (2002) study. However, interviews occurred either immediately after the event or three to four days later. When these shorter delays were used, 3- to 4-year-olds benefitted from training; two of these studies involved only 3- to 4-year-old participants, and both found significant training effects (Thierry et al., 2010; Thierry & Spence, 2002).

A third study by Thierry and colleagues in 2001 included 3- to 4-year-old participants as well as an older comparison group of 5- to 6-year-olds. This study involved exposure to live science demonstrations and video-based demonstrations. Immediately after viewing the presentations, children were asked either source-monitoring questions (training condition) or recognition questions (control condition) about the event. The children then provided free recall reports about the event and finally, a target interview including misleading questions about the sources was conducted. In this study feedback was not provided during the source-monitoring training, but simply answering the source-monitoring questions led to a training effect for the 3- to 4-year-old age group, who provided fewer incorrect responses to misleading questions about source. There were no differences between the 5- to 6-year-olds who participated in the source-monitoring task versus the recognition task. Note that these results are inconsistent with the findings of Poole and Lindsay's (2002) study: whereas Poole and Lindsay had found training effects only for the older children (7- to 8-year-olds), Thierry and colleagues (2001) found training effects only for the younger children (3- to 4-year-olds).

Thierry et al. (2001) reported that 53% of the 5- to 6-year-olds in the control group had spontaneously referred to source during the recognition questions. In addition, the free recall task may have served as a source-monitoring practice for the 5-to 6-year-olds because they were asked to recall information from one source and then the other (i.e., separating their recall by source and drawing attention to the separate sources). It seemed that because the 5- to 6-year-olds were more likely to spontaneously use a strategy without being instructed, the control group was performing similarly to the training group. In a follow-up study where the 5- to 6-year-olds were not given a free recall task, differences between the control and training conditions were evident for the older children (Thierry et al., 2001; Experiment 2). Similarly, Thierry later conducted another study in which both 3- to 4-year-olds and 5- to 6-year-olds benefited from training; however, if supportiveness was increased in the control condition by showing children in both conditions pictures that corresponded with the story and real-life response options during the test, the 5- to 6-year-olds no longer showed training effects (Thierry, 2009).

To summarize, there are two main themes that emerge from the literature on source-monitoring training. The first is that there is evidence that a training procedure that draws attention to source information and encourages strategy use can be effective in helping children to monitor sources more carefully. The training effects in the study by Thierry and colleagues (2001) are particularly notable because no feedback was given after the source-monitoring questions. Being asked source questions was enough to draw the children's attention to the importance of source, creating a training effect (Thierry et al., 2001).

The second theme that emerged was that there are conflicting findings about the trainability of children of different ages. Differences in methodology between studies conducted by Poole and Lindsay (2002) and Thierry and colleagues (2001) likely impacted the difficulty of source-monitoring decisions (i.e., a delay of three months versus no delay), and this may have been a contributing factor in the inconsistent age effects that were observed. Poole and Lindsay (2002) found that older children could be trained, but because the task was very difficult, younger children could not be trained to monitor sources more accurately. Thierry and colleagues (2001) found that when the task was easier young children could be trained, but older children spontaneously “trained” themselves, so training effects were only evident if the control group’s opportunity to produce a strategy and rehearse source was removed.

Comparing results across studies, it seems that when source decisions are very difficult, young children cannot be trained because they will do poorly on the task regardless of having an opportunity to practice. However, older children benefit from the scaffolding effect of training that helps them produce an effective strategy (as in Poole and Lindsay, 2002). When source decisions are easier, older children do not show a training benefit over a control group because children spontaneously produce a strategy and use it effectively regardless of interview condition. However, younger children benefit from the scaffolding effect of training, and when the task is within their developmental norms, they actually have a chance at improving (as in Thierry et al., 2001). No study to date has directly compared the effectiveness of training for different age groups at shorter and longer delays, so incorporating task difficulty into future research is an essential next step for this area, and one that this dissertation addressed.

There were also several other differences between these two studies, including the timing of when the training occurred and the types of questions that were asked (i.e., free recall versus specific questions). Thierry et al. (2001) used a criterion that participants must answer four questions correctly in a row to indicate that they had successfully completed training, whereas Poole and Lindsay (2002) used a set number of questions for training. Thierry et al. used implicit training by simply asking source questions, whereas Poole and Lindsay gave explicit feedback about source information telling children whether they were correct or not. Thierry et al. did the training using target events, but Poole and Lindsay conducted training on non-target events before the target memory interview. In the present research, in order to isolate and manipulate one difference between these studies to examine task difficulty, all other differences were held constant. Therefore, in the source-monitoring training procedures used in the present research, all children were trained on non-target materials, asked direct questions, given feedback about their responses, and trained to a criterion of four questions correct in a row. Children were trained to a criterion and given feedback about their sources decisions in order to maximize the effects of the training procedure. The interview procedure used direct questions so that the memory test was in the same format as the training procedure.

The Present Program of Research

The present studies focussed on questions in two areas of research described in the literature review above: the relationship between executive function and source monitoring, and the effectiveness of source-monitoring training with different age groups, in relation to difficult source-monitoring decisions. With regard to executive function, there are clearly mixed findings about the relationship with source monitoring, with some researchers finding significant relationships and others finding relationships only for certain types of source-monitoring tasks. With such inconsistent findings, it is clear that there must be other factors influencing the strength of these relationships. No study has systematically examined the relationship between executive function and source monitoring with respect to task difficulty. Therefore, the first goal of this dissertation was to explore the relationship between executive function and source monitoring with varying task difficulty and across different types of tasks to illuminate the role that developing executive function skills play in children's increasing ability to monitor sources. By testing this relationship in a variety of conditions including easier and more difficult decisions in Study 1, and different types of decisions (i.e., moving from studying two perceptually distinct sources to a series of repeated events in Study 2), it was possible to find out more about the relationship between these two constructs and the factors that may influence the strength of that relationship.

Existing research on source-monitoring training is inconclusive, because different groups of researchers have found different results regarding which age groups benefit from the training. This is likely due to methodological differences in the way these studies were conducted. One difference among previous studies is the delay that was used; either a short delay (immediate or 3 to 4 days later) leading to an easier source decision, or a long delay (3 months) leading to a more difficult source decision. No research to date has directly compared the effectiveness of the training at shorter and longer delays, but it is quite possible that this could be a major factor that provides insight about the conflicting findings of past research. In addition, source-monitoring training has never been applied to a repeated-event paradigm. Therefore, the second goal of this dissertation was to examine factors that moderate the effectiveness of the source-monitoring training procedure. Studying source-monitoring training in relation to task difficulty in Study 1 and with repeated events in Study 2 created more generalizable conditions including longer delays and more than two sources. If the goal is to generalize these findings to forensic contexts then it is essential that we subject the procedure to rigorous scientific testing in conditions as close as possible to what children would be asked to do in the real world.

To address these research questions, two studies were conducted in which executive function *and* source-monitoring training were explored first in relation to task difficulty in Study

1, and then in the context of a repeated-event paradigm in Study 2. In the first study, task difficulty was manipulated by testing participants' memory after a shorter delay (1-2 days) or a longer delay (8-10 days). Testing children's source monitoring with varying task difficulty may help to explain inconsistencies in the previous literature about the relationship between source monitoring and executive function. This study also examined whether source-monitoring training had different effects for younger versus older children when the task difficulty was varied.

In the second study, children's source monitoring was examined in the context of a repeated-event paradigm, in which source decisions were difficult. The inclusion of more than two sources (i.e., events) created a more realistic and generalizable task. The relationship between source monitoring and executive function was explored under conditions where there were more than two sources and the sources were highly similar. In addition, source-monitoring training has never been examined with repeated events, but it must be carefully tested in this context before recommendations can be made to professionals who interview children about repeated experiences. This study examined whether source-monitoring training benefits source decisions that require differentiating between multiple instances of repeated events.

Study 1: Source Monitoring and Task Difficulty

There are conflicting findings regarding the relationship of executive function with source monitoring in the existing literature, indicating that there may be other factors influencing how the strength of this relationship presents in different studies. One of these factors could potentially be the difficulty of the source-monitoring decisions, as this varies between studies attempting to demonstrate this relationship. Differences in the difficulty of source-monitoring decisions may also be a contributing factor in the inconsistencies in the literature regarding which age groups benefit from source-monitoring training. Therefore, the first study of this dissertation examined task difficulty in the context of both of these research areas; first, how task difficulty influenced the relationship between executive function and source monitoring, and second, how the effectiveness of source-monitoring training varied depending on task difficulty. Task difficulty was manipulated by having children in one condition complete the memory test after a shorter delay (1-2 days) whereas children in the other condition completed the memory test after a longer delay (8-10 days).

Participants were children aged 4-5 or 7-8 years who were exposed to two sources during a single session (a storybook and a real-life science activity), and their encoding was measured immediately afterwards. After either a shorter or longer delay, half of the children were randomly assigned to receive source-monitoring training before a memory interview, whereas the other half of the children did not receive training about sources. All children completed a battery of cognitive tests measuring inhibitory control, working memory, and receptive language. The hypotheses of the study were as follows:

1. It was expected that there would be developmental differences in executive function measures (i.e., working memory and inhibitory control) as well as memory measures (i.e.,

encoding, recognition, and source monitoring),. Specifically, older children (7- to 8-year-olds) were expected to outperform younger children (4- to 5-year-olds) on all measures.

2. It was expected that a positive relationship would be revealed between measures of executive function with both recognition and source monitoring, adding to the body of literature seeking to clarify this relationship.
3. With respect to the executive function-source monitoring relationship, an interaction with delay condition was also expected, such that the relationship would be stronger in the longer delay condition than in the shorter delay condition. Because of the effortful processing required for more difficult source-monitoring decisions, skills like inhibitory control and working memory would be expected to be more necessary for effective source monitoring in the longer delay condition. Such an outcome may explain why some previous studies have found relationships between executive function and source monitoring, while others have not.
4. It was expected that executive function and encoding scores would mediate the relationship between age and source monitoring. That is, these two factors would help to explain how children's source monitoring improves across early childhood; as children get older, their executive function and encoding skills improve, which leads to better source monitoring.
5. In terms of source-monitoring training effects, a three-way interaction between age group, delay condition and training condition was expected such that for younger children, the training would have a larger impact relative to the control group in the shorter delay condition than the longer delay condition, whereas for older children, the training would have a larger impact relative to the control group in the longer delay condition than the shorter delay condition. This hypothesis was grounded in past literature on source-

monitoring training indicating that after short delays training effects were observed only for younger age groups, and after long delays training effects were observed only for older age groups. However, no study has directly compared these effects across short and long delays before. The design of Study 1 may be able to clarify questions in the field regarding the contradictory findings about which age groups benefit from source-monitoring training.

Method

Design

Study 1 used a 2 (Age group: 4-5, 7-8) x 2 (Delay: Shorter, Longer) x 2 (Interview condition: Training, Control) between-subjects factorial design.

Participants

Initially, 234 children who were either 4- to 5 or 7- to 8-years-old were recruited for the study. In total, 44 children were excluded from the study; 9 were absent on the day of the interview, 2 were excluded due to exceptionalities related to autism, 2 were the wrong age for the study, 6 did not understand the interview or could not respond due to language difficulties, 13 had a response bias (e.g., saying “no” to recognition questions about every item during the memory test), and 12 were excluded due to interviewer errors (e.g., mentioning the sources during training for children in the control condition). This left a final sample of 190 children, 47% female, with a mean age of 6.52 years (*SD* = 1.40). See Table 1 for sample sizes within each cell.

The socioeconomic status of participants was estimated by using the highest level of education obtained by the child’s mother as a proxy. Generally the SES was high, with 74% of mothers attending post-secondary education; 33% had attended college, 32% had a Bachelor’s degree, and 9% had a graduate degree. A high school diploma was the highest level of education for 22% of the sample, and 4% chose not to report education levels. Parents reported their children's ethnic background in response to an open-ended prompt, and their responses were later categorized. Ethnicity was not reported for 12% of the sample. Overall, 39% self-identified as Caucasian, 28% identified simply as "Canadian", 7% reported mixed ethnicity, 5% were Asian, and 1% were African American. The remaining 8% self-identified as part of another minority group, and were categorized together because there were so few responses for each (e.g., Metis, Hispanic,.

The children were recruited through the local school board, daycares, summer day camps, and a database of families from the community. Parental consent was obtained prior to participation in the study, and children also gave verbal assent at each session. For children who participated at schools or daycares, compensation was provided through a $10 donation to the school or daycare for each consent form returned. Families from the community who came to the lab to participate were compensated with a total of $20 ($5 per session) to cover travel expenses, and the children also received a small toy at the end of each session.

Materials and Procedure

Each child participated in four sessions with a research assistant. The first session was the to-be-remembered event, which consisted of science presentations from two sources. The second session occurred after a shorter or longer delay period, and comprised a training session and a memory interview about the event. In the third and fourth sessions children completed a battery of cognitive tests.

Event session. Children participated in the event in groups of up to nine children (*M* = 4.91 children per group, *SD* = 2.64). A research assistant presented a live activity and a storybook about the human body, while a second research assistant ensured that the children were paying attention. Both presentations were approximately 10 minutes in length, so the total time for the event was 20 minutes. Each of the presentations contained 12 target details (six props and six actions) that were tested at the interview, for a total of 24 target details. During the event the research assistant verbally emphasized each of the target details to increase the likelihood that the children would encode them (e.g., "Everyone look at these *red apples*. Did you know that *apples* are healthy for you?"). The media of the two presentations was counterbalanced such that each set of activities was shown as the live activity half of the time and the storybook half of the time.

Encoding Assessment. Immediately after the event, children were asked 10 direct questions about the activities to measure their baseline encoding of the event. The questions were about items and actions that were not target details tested at the final memory interview, and were asked in random order across participants. The encoding assessment was conducted to ensure that there were no differences between interview conditions in initial event memory, and also allowed for analyses of the role of encoding in source-monitoring development.

Delay. In the shorter delay condition, the delay between the event and the interview was 1-2 days, whereas in the longer delay condition, the interview occurred 8-10 days after the event. A delay of 8-10 days was chosen for the longer delay condition to ensure a notable delay, but also allow children in the younger age group to remember the event.

Training Phase. After the delay, and immediately prior to the target interview, children individually experienced a live activity and a story about frogs with a different research assistant than the one who conducted the human body event. Each presentation contained six target details (three props and three actions), for a total of 12 target details. The story and live activity were the same length, and this event took approximately 5 min. Again, the scripts were counterbalanced so that each presentation script was shown as the story half of the time and the live activity half of the time. After the frog event children were given a memory test of up to 18 questions (12 about the target details and 6 about details not presented at the activities). The procedure for this test differed depending upon whether children were in the control condition or the training condition.

In the control group, children were asked recognition questions about the frog activities and were provided with feedback about their answers (e.g., "You're right! You did learn that frogs have bumpy backs!"). The interviewer continued asking recognition questions until they met the criterion of four consecutive correct responses, and then the interviewer proceeded to the target interview. The sources were never mentioned to the children during the training phase in the control condition. In the training condition, children were asked recognition *and source* questions and were given feedback about their answers (e.g., "Actually you didn't learn that in the story, you learned that in the real-life activity"). Children in this condition were asked recognition and source question pairs until they answered four consecutive question pairs correctly (i.e., they answered both the recognition and source questions correctly), and then the interviewer proceeded to the target interview.

Target Interview. Immediately after the training session, children completed the target interview about the human body activities. The interview was an oral forced choice test with 36 questions pairs. See Appendix A for a list of questions. These questions represented the 12 target details from each source during the human body activities, as well as 12 misleading details that were not present at the activities. For each detail children were asked recognition and source questions (e.g., "Were there goldfish crackers at the human body activities?" If yes, "Were the goldfish crackers in the story or the real-life activity?"). Children did not receive feedback on these responses, as the interviewers were blind to counterbalancing conditions and as such, did not know which source was correct for any of the details. The questions in the target interview were asked in random order. The interviewer alternated the phrasing of source choices such that half of the time she asked whether a detail occurred “in the story or the real-life activity”, and half of the time she asked whether a detail occurred “in the real-life activity or the story”. This was done to prevent response biases.

Cognitive Assessments. Within approximately one week of the interview, children completed a third and fourth session which consisted of a battery of cognitive tests. These tests were administered across two sessions to reduce participant fatigue, as the total time for the battery of tests was approximately 45 minutes. These assessments included two working memory measures, an inhibitory control measure, and a receptive language test. In most cases, children completed the receptive language test in one session of up to 25 minutes, and then completed all other tests at a separate session of up to 20 minutes.

WISC Digit Span. The WISC-IV Digit Span subtest was used to measure working memory (Wechsler, 2003), and this was described to the children as a number game. For the Forward Digit Span participants repeated a sequence of digits that was read out by the administrator. The number sequences got progressively longer, beginning with two digits and potentially progressing up to a series of nine digits. Testing continued until the child had failed two trials of a sequence length. The Backward Digit Span test was conducted in a similar way, but participants heard a sequence of numbers and had to repeat the numbers in the reverse order. Participants were given an example with two digits and one practice trial with feedback before testing began. If the participants did not answer the practice trial correctly, they were given up to two more practice trials, at which point testing was discontinued if they still did not understand.

The test began with two digits and the sequences increased in length up to eight digits. Once a child had tried and failed twice with a sequence length, testing ceased.

One point was scored for each time children correctly repeated a series of numbers (or repeated the numbers in backwards order, for the backward digit span test). The forward and backward digit span scores were totalled, and scores could range from 0 to 30.The WISC-IV Digit Span has been shown to have an internal consistency reliability of .87, and test-retest reliability of .81 (Williams, Weiss & Rolfhus, 2003).

Pick the Picture. The second measure of working memory was a Pick the Picture task (Willoughby, Wirth & Blair, 2012). Children were presented with a set of pictures and were asked to pick each picture once so that all of the pictures would "get a turn". That is, children should never choose a picture that they have chosen before. This task requires working memory because children must remember which pictures they have already chosen in order to avoid choosing the same picture twice. For example, in the first set there was an apple and a dog. Children could choose either of the two pictures. On the second page, the apple and the dog appeared again and children were asked to choose a different picture than what they had already chosen. The task began with a set of two pictures, and then progressed to sets of three, four, six and eight pictures. Children were given two sets with each number of pictures. The pictures in each set were new pictures that were not seen in a previous set. The spatial location of the pictures in each set changed on each page so that children could not use location to infer which pictures they had chosen before. Previous research indicated that use of up to six pictures in a set was appropriate for 4- to 6-year-olds. Because we included older children in our sample, we added the eight-picture set.

One point was scored for each correct selection of an item that was not previously chosen. Given that the first picture chosen in each of the 10 sets was automatically correct, scores could range from 10 to 44. Willoughby, Blair, Wirth and Greenberg (2012) assessed the criterion validity of this measure as well as how reliability differed as a function of ability level. They reported reliability estimates greater than .70 for children who were between two standard deviations below the mean to one standard deviation above the mean. Scores on this measure were also significantly correlated with five measures of academic achievement.

Simon Task. The inhibitory control task was a Simon task (Roebers & Kauer, 2009; Willoughby, Wirth, Blair & Greenberg, 2010; Gerardi-Caulton, 2000) that was presented on a laptop computer using Superlab**.** This task involved a blue and a yellow starfish that appeared at either the left or the right side of the computer screen (see Appendix B for stimuli). On the left side of the keyboard was a blue button, and on the right side of the keyboard was a yellow button. Children were instructed to push the blue button if they saw a blue starfish, and the yellow button if they saw a yellow starfish. If the blue starfish was on the same side of the computer as the blue button, this was a congruent trial. If the blue starfish was on the opposite side of the computer from the blue button, it was an incongruent trial (and the same for the yellow starfish). Therefore, there were four types of trials (blue congruent, blue incongruent, yellow congruent, yellow incongruent).

The pictures were visible until the children responded, and all trials were separated by a central fixation cross lasting 1 second. Participants were encouraged to respond as quickly as they could, without making any mistakes. To ensure that the children understood the task, they were given four congruent practice trials with feedback indicating whether they were correct or incorrect. If they did not get at least three of these trials correct, they were given a new set of practice trials before proceeding to the test phase. The task consisted of 40 randomized trials, 10 of each type. After 20 trials children took a short break and were given a sticker and positive feedback, and then were reminded of the instructions before continuing with the last 20 trials.

Correct and incorrect responses were recorded using Superlab software. However, in previous research in the same lab (Earhart & Roberts, 2014), a similar inhibitory control task scored for the accuracy of responses showed evidence of ceiling effects amongst 7- to 8-year-old children. It was expected that there might be greater variability in reaction times than there would be for a measure of correct judgments alone because even if all children performed well on the task, older children may complete the trials more quickly than younger children (Williams, Ponesse, Schachar, Logan & Tannock, 1999). Faster reaction times would indicate less of a response cost associated with inhibiting a habitual response, and therefore, better inhibitory control. Thus, reaction times in milliseconds were recorded as well.

Accuracy was coded by assigning one point for each correct trial (i.e., pressing the correct coloured button to match the starfish, regardless of its location). Scores could range from 0 to 40. Incorrect trials were not included in the calculation of average reaction times because children were not successfully inhibiting their behavioural responses on these trials. Reaction times less than 100 ms were considered premature, and were not included. As well, any reaction times over 10 seconds were excluded because the children were presumably distracted. The remaining reaction times were used to compute an average response time for congruent and incongruent trials for each participant. Scores for each participant were then computed by subtracting their average reaction times on congruent trials from their average reactions times on incongruent trials, yielding a score that represented the response cost associated with inhibition corrected for general reaction time (termed the "Simon Effect" in previous literature; e.g., Martin-Rhee & Bialystok, 2008).

Peabody Picture Vocabulary Test. Children completed the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007), which is a standardized test of receptive vocabulary. In this test, children were shown four pictures while the administrator read a word and the children were instructed to point to the picture that matched the word. Testing began at an age-appropriate baseline level and continued through progressively more challenging levels until the children made eight or more mistakes in one level, at which point testing was discontinued. The PPVT was scored with one point for each picture that the children correctly identified. Because the test has different starting points for children of different ages, a base number was added to the final score which assumed correct identification of all words below the level where they started.

This test is appropriate for age 2.5 years through adulthood. The internal consistency reliability is .94, and the test-retest reliability is .93 (Dunn & Dunn, 2007).

Interview Accuracy Coding

Encoding Assessment. The encoding assessment that was administered immediately after the event was scored with one point for each correct answer, and thus, scores could range from 0 to 10. Some children were not asked all 10 questions due to time constraints or interviewer errors, so proportion scores were computed by dividing the number of correct responses by the number of questions asked.

Recognition Accuracy. Recognition questions from the target interview were scored by calculating proportion scores for non-misleading recognition (correct identification of details that were present at the activities) and false alarms (incorrect identification of details that were not part of the activities as having been present). Recognition accuracy scores were then calculated by subtracting the proportion of false alarms from the proportion of non-misleading recognition for each participant, and thus, scores could range from -1 to 1. “Don’t know” responses were conservatively coded as incorrect.

Source Accuracy. During the target interview, children were only asked source questions about the items that they recognized from the activities, so children were asked different numbers of source questions depending on their recognition responses. Therefore, proportion scores were calculated by dividing the number of correct source attributions by the number of source questions asked.

Inter-rater Reliability. The nature of the coding was objective (e.g., children’s yes/no responses were either correct or incorrect based on the details at the activities), so interrater reliability was greater than 99%. All of the data were coded by two raters to ensure accuracy. The few disagreements were due to addition errors and were resolved before data analysis.

Results

Analytic Strategy

Preliminary analyses were first conducted to examine the distributions and factor structure of the executive function scores. Preliminary analyses of the memory scores tested whether the event leader, interviewer, event order, or gender were related to recognition or source memory. Age, receptive language, executive function and initial event encoding were also tested to verify that there were no differences between conditions. Next, developmental differences in the executive function scores and the recognition and source accuracy scores were analyzed. The relationships between executive function scores and recognition and source memory were tested, and working memory and encoding were explored as mediators between age and source monitoring. Finally, training and delay conditions were examined in relation to children's source-monitoring accuracy. An alpha level of .05 was used to determine significance for all analyses.

Preliminary Analyses

Executive Function. ***Descriptives*****.** Inhibitory control scores were first examined. Accuracy scores for the Simon task showed evidence of a ceiling effect, with a negatively skewed distribution and little variability (see Figure 1). The mean score was 35.59 (*SD* = 4.38) and the median was 37 out of 40. The reaction times, however, showed more variability with an approximately normal distribution (see Figure 2). Therefore, reaction times were used in further analyses of inhibitory control rather than accuracy scores. The average difference between congruent and incongruent trials was approximately 200 ms, with a range from -991ms (with a negative score indicating faster average performance on incongruent trials) to 1271ms.

The working memory scores were examined next. The Pick the Picture data showed a similar ceiling effect to the Simon task accuracy scores, with a mean of 41.69 (*SD* = 2.51) and a median score of 42 out of 44; generally children made very few errors on this task (see Figure 3). The WISC scores, on the other hand, showed more variability and no skew in the data, with a relatively normal distribution (see Figure 4). Scores ranged from 3 to 22, with a mean of 11.44 (*SD* = 3.23).

Factor Analysis**.** A factor analysis of the scores from the 3 executive function tasks (WISC, Pick the Picture, and Simon task reaction times) revealed two factors, which explained 81% of the variance. The two working memory scores loaded heavily on the first factor and the inhibitory control scores loaded on the other factor (see Table 2 for factor loadings). This indicates that these tasks differentiated between the two aspects of executive function that they were intended to measure.

In addition to loading on the same factor, the working memory scores were significantly correlated with each other, r (174) = .44, p <.001. However, composite scores were not considered appropriate due to concerns that the lack of variability in the Pick the Picture scores would make it difficult to find significant effects. This could potentially mask the true relationship of working memory with source monitoring that could be detected with the more psychometrically sound WISC scores separately. Therefore, all three executive function scores were entered separately in all further analyses of executive function.

All multiple regression analyses were also run without the Pick the Picture scores. Some of the results differed based on whether the Pick the Picture scores were entered or not (as noted below, where applicable), and thus, because of the exploratory nature of the dissertation, the Pick the Picture scores were included despite the restricted range of scores.

Recognition and Source-Monitoring Scores. Three methodological variables were tested to ensure that they were not systematically related to either recognition or source memory. Four one-way analyses of variance (ANOVAs) were conducted with the event leader (one of five female research assistants), and the interviewer (one of five female research assistants) as the independent variables and recognition and source accuracy as the dependent variables. All were non-significant, Fs $\leq$ 2.26, ps $\geq$ .07, $\eta^2_p \leq .05$, indicating that the event leader and interviewer had no impact on either recognition or source accuracy. Two independent-samples t-tests confirmed no effect of the order of the two presentations (real-life or story first) on recognition or source accuracy, ts $\leq |1.54|$, ps $\geq$ .13, Cohen's d = 0.22. There were also no gender differences in either recognition or source accuracy, as assessed by two independent-samples t-tests, ts $\leq |0.29|$, ps $\geq$ .77, Cohen's ds $\leq$ 0.04. Therefore, there was no need to covary these variables in any further analyses.

Although children were randomly assigned to conditions and any other characteristics should statistically even out between groups, preliminary analyses were conducted to ensure that there were no differences between the children assigned to the control and training conditions, or the shorter and longer delay conditions. A series of independent samples *t*-tests comparing the control and training conditions found no significant differences in receptive language, executive function measures, age, delay between event and interview, or initial event encoding, $t\text{s} \leq |1.35|$, $p\text{s} \geq .18$, Cohen's $d\text{s} \leq 0.23$. A series of independent samples *t*-tests comparing the shorter and longer delay conditions found no significant differences in receptive language, executive function measures, age, or initial event encoding, $t\text{s} \leq |1.49|$, $p\text{s} \geq .14$, Cohen's $d\text{s} \leq 0.23$. An independent samples *t*-test showed no significant difference between the younger and older age groups in the average delay between event and interview, $t\ (188) = 0.10$, $p = .92$, Cohen's $d = 0.01$.

Training. Of the 190 children in the sample, 155 passed the criterion of four consecutive correct questions/question pairs before proceeding to the target interview. Thirty-five participants were asked all 18 questions/question pairs but never met the criterion before proceeding to the target interview. Of the children that did not meet the criterion, 31 were younger children, whereas only 4 were older children, and 12 were in the control condition, whereas the other 23 were in the training condition. For those that did meet the criterion, it took six questions on average to get four correct in a row ($M = 6.34$, $SD = 3.09$).

Removing children who did not pass the criterion (the majority of whom were in the training condition) might artificially inflate the mean for the training group because the children with the least well-developed source-monitoring skills were not included in analyses. In addition, excluding the children who did not pass criterion would significantly reduce the sample sizes in some cells, particularly for the younger children. Therefore, it was conservatively decided that all participants who completed the procedure would be included in further analyses, regardless of whether or not they met the criterion.

***Descriptives*.** For the overall sample, children's initial event encoding proportion scores ranged from 0 to 1.00; the mean was .65 (*SD* = .21). Generally, the recognition and source accuracy scores in the target interview were quite high for most children. Recognition accuracy scores ranged from -.08 to .92, with a mean of .53 (*SD* = .19). Source accuracy proportion scores ranged from .25 to 1.00, with a mean of .77 (*SD* = .14).

A paired-samples *t*-test compared children's recognition accuracy proportion scores for non-misleading versus misleading recognition (i.e., identifying items that were present during the activities, versus rejecting items that were not present). On average, children's accuracy was significantly higher for rejecting distractor items ($M = .82$, $SD = .17$) than for correctly recognizing target details ($M = .68$, $SD = .15$; t [189] = -7.79, $p < .001$, Cohen's $d = 0.87$), indicating that the most common error was forgetting things that had happened during the science activities.

Two paired-samples *t*-tests compared accuracy rates for the story and the real-life presentation, for both recognition and source judgments. Both were significant, ts ≥ 10.40, ps < .001, Cohen's ds ≥ 1.51. Children recognized more items from the real-life presentation than the story, and correctly attributed the source of items from the real-life presentation more often.

A 2 (Age: 4-5, 7-8) x 2 (Delay: Shorter, Longer) between-subjects ANOVA assessed delay condition differences on children's recognition accuracy. There were main effects of Age, Delay, and an Age by Delay interaction (see Table 3 for test statistics). The Age by Delay interaction was followed up with post-hoc *t*-tests examining the effect of the delay on younger and older children's recognition separately. Delay was significant for both age groups, *t*s $\geq$ 2.54, *p*s $\leq$ .01, Cohen's *d*s $\geq$ 0.53, but had a larger effect on older children's recognition accuracy; after a longer delay the younger children's recognition accuracy dropped from .53 (*SD* = .18) to .44 (*SD* = .17), and older children's recognition accuracy dropped from .67 (*SD* = .15) to .49 (*SD* = .15).

Inferential Analyses

Developmental Differences. Age was treated as a continuous variable in all regression analyses. It was hypothesized that age would be related to both the executive function and memory variables. To test for relationships with executive fulknction, age was entered as the independent variable in three linear regression analyses using the inhibitory control and working memory measures as dependent variables. Age was not significantly related to the inhibitory control scores, F (1, 143) = 1.09, p = .30, R^2 = .01, but it was related to scores on both working memory measures, *F*s > 37.97, *p*s <.001. $R^2 \geq .18$.

To examine the relationships between age and encoding and memory accuracy, three linear regressions were conducted with age as the independent variable and encoding, recognition, and source memory as the dependent variables. Age was significant in all three analyses, *F*s > 3.93, *p*s < .05, $R^2 \geq .02$. See Table 4 for relevant statistics for significant regression analyses.

The Relationship between Executive Function and Memory Performance. It was hypothesized that both working memory and inhibitory control would be significantly related to recognition and source-monitoring accuracy. A multiple regression analysis was conducted using working memory and inhibitory control scores as predictors of recognition memory. The model was significant, F (3, 137) = 10.87, p < .001, R^2 = .19. WISC scores were significantly related to recognition accuracy, but the Simon task scores did not reach traditional levels of significance (*p* = .09). Pick the Picture scores were not related to recognition accuracy. A second regression was run with the same independent variables, but also controlling for age, and the results were the same. See Table 5 for test statistics associated with each independent variable in the regressions. These analyses revealed that working memory was related to recognition, even when controlling for age, whereas inhibitory control showed a non-significant trend towards a relationship with recognition accuracy.

A multiple regression was run with the three executive function measures as independent variables and source-monitoring accuracy as the dependent variable. Again, the model was significant, F (3, 137) = 5.78, $p = .001$, $R^2 = .11$. The WISC and Simon scores were both significantly related to source-monitoring scores, but Pick the Picture scores did not reach traditional levels of significance ($p = .09$)[1]. When the analysis was run with a step-wise variable selection technique, the WISC scores were entered first with an $R^2 = .07$. The R^2 change when the Simon scores were entered was .04, indicating that working memory scores accounted for more variance in source-monitoring scores than inhibitory control scores did. Another regression was conducted with the same independent variables but this time also controlling for age, and the results were the same. See Table 5 for relevant test statistics.

It was also hypothesized that the relationships between executive function and source-monitoring accuracy would be related to the delay condition, such that the relationships would be stronger in the longer delay condition. To examine whether the relationships between executive function scores and source monitoring were influenced by task difficulty, three interaction terms

---

1 When this analysis was re-run without the Pick the Picture scores as an independent variable, the results were slightly different. WISC scores were still significant, but the Simon scores no longer reached traditional levels of significance, t (139) = -1.85, $p = .07$.

were created to assess an interaction between delay condition and the three executive function scores on the dependent variable. The model was significant, F (7, 133) = 3.12, p = .004, R^2 = .14, but none of the predictors were, indicating a problem with multicollinearity. The executive function variables were standardized to correct this problem, and the analysis was conducted again. None of the interaction terms with delay were significant, ts ≤ 1.61, ps ≥ .11. Thus, the relationships between executive function and source monitoring were not related to the delay condition.

Mediation of the Age-Source Monitoring Relationship. It was hypothesized that executive function and encoding could potentially mediate the relationship between age and source-monitoring accuracy. However, because the inhibitory control scores were not related to age, they were not considered as a mediator. Similarly, Pick the Picture scores were not significantly related to source-monitoring accuracy, and therefore these scores were also not considered as a mediator. WISC working memory scores and encoding were considered to test the possibility that they could explain the developmental progression in source-monitoring accuracy.

Baron and Kenny’s (1986) four-step method for testing mediation was used. This method uses a series of regressions to show that a) the independent variable predicts the dependent variable, b) the independent variable predicts the mediator, c) the mediator predicts the dependent variable, and d) when the independent variable and mediator are entered together to predict the dependent variable, only the mediator is significant.

Regression analyses reported above indicated that age was related to source-monitoring accuracy, and also made an independent contribution in predicting WISC scores. An additional regression analysis was run to determine whether WISC scores were related to source-monitoring accuracy, and it was significant. However, when age and WISC scores were entered together, only the WISC scores remained significant (see Figure 5; test statistics and significance levels at each step are listed in Table 6). This demonstrated that working memory mediated the relationship between age and source monitoring, such that as children get older, their working memory improves, which leads to improved source monitoring.

Similarly, regression analyses reported above indicated that age was related to source-monitoring accuracy, and also independently predicted encoding scores. An additional regression analysis was run to determine whether encoding scores were also related to source-monitoring accuracy, and this analysis revealed a significant relationship. When age and encoding scores were entered simultaneously, only the encoding scores were significant (see Figure 6; test statistics and significance levels at each step are listed in Table 7). Initial encoding of the event mediated the developmental progression in source monitoring, indicating that as children get older, they encode events more accurately, which leads to better source-monitoring accuracy.

Given that both working memory and encoding were shown to mediate age differences in source monitoring, a further exploratory analysis examined whether improved encoding could explain the role of working memory in source monitoring. Improved working memory could lead to better encoding because children would be better able to bind the features of events together during the encoding process, making source monitoring decisions more accurate (Mammarella & Fairfield, 2007). Working memory was already shown to be related to source monitoring, as was encoding, as reported in the two paragraphs above. A further regression analysis showed that working memory was related to event encoding, but when working memory and encoding were entered together, working memory was no longer significantly related to source monitoring (see Figure 7; test statistics and significance levels at each step are listed in Table 8). Together, these results demonstrate a double mediation in which as children get older, their working memory improves, leading to stronger event encoding, which in turn leads to more accurate source monitoring.

Source-Monitoring Training Effects. A three-way interaction between age group, interview condition and delay condition on source-monitoring accuracy was predicted, such that younger children would benefit more from the training in the shorter delay condition, whereas older children would benefit more from the training in the longer delay condition. A 2 (Age: 4-5, 7-8) x 2 (Interview condition: Control, Training) x 2 (Delay: Shorter, Longer) between-subjects ANOVA with source accuracy as the dependent variable assessed training effects on source monitoring, and whether they were influenced by age group or delay condition. There were main effects of Delay and Training, a marginal main effect of Age Group, and a three-way interaction (see Table 9 for test statistics for significant effects). When the same analysis was run using encoding scores as a covariate the results were the same, except that there was no longer a marginal main effect of age group.

Two follow-up 2 (Delay: Shorter, Longer) x 2 (Interview condition: Training, Control) ANOVAs were run for younger and older children separately to examine the three-way interaction further. For younger children, there was only a main effect of Delay, F (1, 91) = 5.76, p = .02, η^2_p = .06, with children in the shorter delay condition (M = .79, SD = .12) outperforming children in the longer delay condition (M = .72, SD = .16). There was no main effect or interaction with training, Fs $\leq$ 1.06, ps $\geq$ .31, η^2_p $\leq$ .01. For older children, there was a main effect of Training, F (1, 91) = 4.95, p = .03, η^2_p = .05, and a Training x Delay interaction, F (1, 91) = 4.72, p = .03, η^2_p = .05. The interaction was followed up with post-hoc t-tests comparing training effects after shorter and longer delays. In the shorter delay condition, older children showed a significant effect of training, t (29.83) = -3.05, p = .01, Cohen's d = 0.90. The children in the training condition (M = .87, SD = .07) had higher source accuracy than those in the control condition (M = .75, SD = .17). In the longer delay condition, the effect of training was not significant, t (45) = -0.38, p = .97, Cohen's d = 0.01. The training and control groups did not differ in their source-monitoring accuracy ($M_{training}$ = .78, SD = .13; $M_{control}$ = .77, SD = .11). Therefore, the only group to benefit from source-monitoring training was the older children in the shorter delay condition.

Discussion

The purpose of this study was two-fold: to examine the relationship between executive function and source monitoring when task difficulty was varied, as well as to measure the effectiveness of source-monitoring training for different age groups when task difficulty was varied. Participants were 190 children aged 4- to 5 or 7- to 8-years-old who experienced an event with two sources. After a shorter (1-2 days) or longer (8-10 days) delay, participants were randomly assigned to receive source-monitoring training or recognition questions, before completing a memory interview about the event. Children also completed measures of working memory, inhibitory control, and receptive language. This study was the first to examine how task difficulty interacts with the effects of source-monitoring training or alters the presentation of the executive function-source monitoring relationship.

Developmental Differences

The present study involved children between the ages of 4- and 8-years-old, and developmental differences in both executive function and memory were expected. Although age was significantly related to working memory scores, surprisingly, age was not related to inhibitory control. This is inconsistent with a large-scale developmental study of reaction times to a Go/No-Go inhibitory control task across the lifespan, which showed decreasing reaction times across childhood (Williams et al., 1999). This may have been a result of the way the reaction times were coded; for example, removing reaction times of over 10 seconds would be likely to affect the scores of younger children more than older children, thus reducing the strength of the relationship. Regardless of the fact that inhibitory control scores were not related to age, individual differences in inhibitory control were related to other variables of interest (such as source-monitoring scores). Thus, the scores seemed to be a meaningful measure of inhibition despite the fact that they were not related to age.

With respect to the memory scores, there were robust developmental differences across the various measures including encoding, recognition, and source memory, consistent with the extensive previous literature (e.g., Roberts, 2002; Lindsay, Johnson & Kwon, 1991). It was important to confirm these age differences in our own data because developmental differences in memory are widely accepted and if they were not found, it may raise methodological concerns that call into question the validity of the data. Interestingly, age interacted with delay condition to influence recognition accuracy scores. The longer delay had a larger detrimental impact on older children's scores than it did on younger children's scores; although there were clear age differences after the shorter delay, the differences were less pronounced after the longer delay. There were also developmental differences in children's abilities to pass the criterion of four consecutive correct question pairs during the training phase; the large majority of the children who did not pass the criterion were younger children. This demonstrates that the source decisions in the training phase were difficult, especially for young children. However, even the children who did not pass the criterion were exposed to the manipulation and received the instructions and materials. Therefore, we included all children in the analyses.

The Relationship between Executive Function and Memory Performance

This study allowed for an examination of how executive function relates to recognition accuracy and source accuracy separately, unlike some other studies which have combined recognition and source tests in their methodology (i.e., by asking one combined recognition and source question such as, "Did you see X in the real life activity, story or neither?" e.g., Roberts & Powell, 2005b; Karpinski & Scullin, 2009). It was expected that executive function would be related to both dependent variables.

For recognition accuracy, the results were mixed; while working memory was clearly related, inhibitory control showed a non-significant trend. This may indicate that some components of executive function are more strongly related to episodic memory than others. For source monitoring, on the other hand, both working memory and inhibitory control were significantly related to source monitoring, providing clear evidence for the role of executive function in source monitoring. Notably, these relationships remained significant even when controlling for age. Thus, these specific cognitive skills were more strongly related to source-monitoring accuracy than the global developmental factor represented by age, which includes developmental changes in many areas such as language, problem solving, and reasoning. At the present time, the research literature examining the relationship between executive function and source monitoring is incomplete, but the present findings add to the growing body of support for the role of executive function, and especially working memory, in both source monitoring and recognition accuracy.

A key hypothesis of the present research was that the relationship of executive function to source monitoring would depend on the difficulty of the task, which was manipulated by using shorter or longer delay periods. Contrary to this hypothesis, there were no interactions between delay condition and any of the executive function measures on source-monitoring performance; the relationship did not present differently depending on the difficulty of the source-monitoring decisions.

It is possible that the delay conditions were not different enough, and that with more drastic differences in task difficulty the relationships would vary. However, this seems unlikely because the shorter and longer delay conditions did affect other dependent variables, producing main effects of delay condition on both recognition and source accuracy. It is also possible that executive function plays a role in source monitoring regardless of the task difficulty, which would mean that task difficulty is not a factor in the diverse results of previous research, and other differences in methodology may help explain these contradictions better. Future research should continue to explore this question by manipulating task difficulty in different ways, as well as continuing to explore other factors that may contribute to the differences in findings of previous studies.

The most interesting findings with regard to executive function and source monitoring were the mediation models that were tested, which can help to explain how source monitoring improves with age. The mediation model linking working memory to age-related improvements in source monitoring is the most conclusive evidence to date that executive function contributes to developmental changes in source monitoring. These findings also add another piece to the puzzle by explaining how children’s developing executive function skills exert their influence on source monitoring; specifically, that this pathway occurs through stronger event encoding. This is consistent with previous research that has emphasized the role of feature binding in source monitoring. In other words, binding the features of an event together at encoding, such as perceptual details and semantic meaning, would lead to a stronger ability to later monitor the sources of those memories (Mammarella & Fairfield, 2007; Kovacs & Newcombe, 2006; Lloyd, Doydum & Newcombe, 2009). This study is the first to demonstrate a developmental pathway that explains age-related changes in source monitoring because as children get older their working memory improves, which leads to better event encoding, and therefore more accurate source monitoring.

Source-Monitoring Training Effects

The final hypothesis of the study was that younger children would benefit more from the training at shorter delays, whereas older children would benefit more from the training at longer delays. These hypotheses were generated on the basis of previous literature showing that when a short delay (immediate or 1 day) was used, younger children benefitted from the source training but older children did not (Thierry et al., 2001). On the other hand, when a long delay (3 months) was used, older children benefitted from the training, whereas younger children did not (Poole & Lindsay, 2002). These findings lend themselves to the explanation that shorter delays are easier source-monitoring tasks, and therefore older children do not need training, whereas longer delays are more difficult source-monitoring tasks and while older children would benefit from training, it is beyond the scope of younger children’s developmental abilities to source monitor under these more challenging conditions. Contrary to this hypothesis, the three-way interaction of age group, delay condition, and interview condition revealed that there were no training effects for the younger children, and that for the older children, only the shorter delay condition benefitted from the training.

Although the hypothesis was not supported, these findings are not completely unexpected, as it is typical in the previous literature to find training effects that are qualified by complex interactions (e.g., only for certain age groups, or only for certain types of source-monitoring tasks). As noted in the introduction, there were many other differences between these two studies aside from the delay period used. For example, Thierry et al. (2001) used implicit training by simply asking source questions, and did the training using target events, whereas Poole and Lindsay (2002) gave explicit feedback during the source training and conducted the training on non-target events before the target memory interview. Because the present study adopted a methodology more similar to Poole and Lindsay's (i.e., using explicit feedback and non-target training events) it is not completely surprising that our findings were similar; Poole and Lindsay (2002) found that only the older children in their sample benefitted from the training, as did the present study. Thierry et al. (2001) found training benefits for younger children using a methodology quite different to that of the current research (i.e., without feedback, and training was conducted on the target events). Overall this indicates that perhaps the varying delay periods did not play a role in the conflicting findings of previous studies of source-monitoring training, but rather other methodological differences may help to explain these findings further. The source-monitoring training effects are clearly contextually sensitive, so future research can continue to isolate and test methodological factors that may help to explain why the previous literature is inconsistent, and may also provide further information about the conditions in which source-monitoring training will be useful.

Because there were no training effects for either age group after a longer delay in the present study, there is doubt about whether the training procedure would be useful in more realistic practical settings, such as in forensic investigations with children, where interviews may take place long after the to-be-remembered event. This study was not designed to test the implementation in real-world settings; rather, the appropriate initial step was to assess the exact nature of these relationships. These findings do, however, highlight the need for further research that examines the effectiveness of source training in more realistic conditions, such as after longer delays and with a series of repeated events, rather than two perceptually distinct sources.

Factor Analysis of Executive Function Scores

The fact that working memory and inhibitory control measures were distinct in a factor analysis of executive function scores provides more information about the structure of executive function. Indeed, there is debate about how many components executive function has, and what they may be (Wiebe et al., 2011; cf. Miller, Giesbrecht, Muller, McInerney & Kerns, 2012). The distinct factor loadings on these measures were surprising because theoretically, one might expect that there would be overlap in what these tasks are measuring. For example, though the Simon task is designed to measure inhibitory control, working memory would be required to remember the rules throughout the task. Despite that fact, there was a strong separation between these measures, which made it possible to determine the specific components of executive function that are or are not related to memory measures. As was noted above, differentiated patterns did arise for the separate components of executive function. For example, working memory predicted recognition scores whereas inhibitory control did not. The clear separation of these measures through the factor analysis allows for confidence in these analyses at the level of individual components of executive function, rather than assessing a global executive function factor in relation to source monitoring.

Distributions of Executive Function Scores

Although the Simon task reaction times and WISC digit span scores had approximately normal distributions with no evidence of ceiling effects, the Pick the Picture scores were skewed with little variability. Analyses were conducted with the WISC and Pick the Picture scores entered separately to avoid masking effects of working memory due to the reduced variability on the Pick the Picture task. Indeed, the Pick the Picture scores were only marginally related to source monitoring, and were not related to recognition accuracy at all, whereas the WISC scores significantly predicted both recognition and source memory. Because there were known problems with the psychometrics of the Pick the Picture task, it was assumed that these non-significant results reflected those limitations, rather than a genuine challenge to the relationship between working memory and the recognition and source accuracy scores.

The ceiling effects found for both Pick the Picture and Simon task *accuracy* scores highlight the difficulty of creating behavioural executive function measures that are appropriate for a wide age range, and yet challenging enough for all age groups to show the variability required for statistical analysis. There are vast differences in the cognitive development of 4- to 8-year-olds, and the problem of finding appropriate measures limits the ability to assess the complex relationship between executive function and source monitoring. Finding developmentally sensitive measures is key to further research in this area, and this study provides evidence that modifying the Pick the Picture task to make it more challenging, as in the present study, was not sufficient to yield an appropriate measure. The Pick the Picture task had been used successfully in previous research with children from 4- to 6-years-old (Willoughby et al., 2012) and it was expected that by extending the number of pictures in each set, the task would be challenging enough for children up to the age of 8. Unfortunately, this was not the case. It is worth considering the role that psychometrics may play in the mixed evidence of the relationship between executive function and source monitoring in the larger literature, considering that in this study, measuring working memory with two different tasks led to different outcomes on each.

Presentation Modality

Both recognition and source judgments were more accurate for target details presented in the real life presentation than those presented in the story. This demonstrates that children tended to remember directly experienced items with perceptual detail better than things that they had only heard about, consistent with the educational literature on the benefits of active/experiential learning (e.g., see Kolb, Boyatzis & Mainemelis, 2001, for a review of Experiential Learning Theory). The items that were rich in perceptual detail were also easier to make source judgments about, as the perceptual detail would indicate that they had been watched in the real life presentation. The fact that children had lower source accuracy for items from the story means that they erred by attributing story items as having been presented in the real-life presentation. This demonstrates a bias towards reasoning that if they remembered an item, it must have been in the real life presentation.

Limitations and Future Directions

One of the limitations of this study, as discussed above, is the issue of finding developmentally appropriate behavioural measures of executive function. Fortunately, some of the measures had good variability across age groups, allowing for meaningful analyses of both working memory and inhibitory control. Perhaps using a different method, such as teacher- or parent-reports of executive function, would limit the ceiling effects seen with behavioural measures and demonstrate the relationships between executive function and memory more clearly. In the second study of this dissertation, the relationship between executive function and source monitoring was measured using parent reports of executive function to explore this possibility further.

Another limitation of the present research is that the procedure was not very generalizable to the conditions in which everyday remembering occurs. Even children in the longer delay condition recalled the event after only 8 to 10 days, but typically when people make source-monitoring decisions in everyday life the delay could be much longer, and in the case of forensic investigations, interviews often happen long after the events in question. As noted earlier, it is important to consider whether the findings would hold true after much longer and more realistic delay periods, and this is an area that future research can address.

The event was unique and novel, consisting of two distinct sources that differed in perceptual detail. In everyday remembering sources may not be this memorable and distinct from one another, which would make source decisions much more difficult. In investigations of child abuse, the sources are often a series of repeated events that children must distinguish between in order to provide details particular to each separate occurrence. This represents a more difficult task as the sources would be highly similar, and there would be more than two sources involved in decision-making.

The second study of this dissertation addressed this problem by using a repeated-event paradigm where children were exposed to a series of similar events and had to describe the details specific to one instance. Both the source-monitoring training literature and the executive function literature could be enhanced by an examination of these effects under more realistic and difficult conditions, which increase the generalizability of the findings. Because there are more than two sources and the sources are highly similar, these decisions are inherently more difficult than the classic studies involving two perceptually distinct sources. Source-monitoring training has never been studied with repeated events before, so it is essential to subject the training procedure to rigorous scientific testing in more generalizable but controlled conditions.

Study 2: Source Monitoring with Repeated Events

Empirical research on source-monitoring processes has typically included only two sources for comparison in the context of a single event. However, it is more realistic (and potentially more interesting) to examine source decisions that involve more than two sources. This would make research on both the cognitive underpinnings of source monitoring and the effectiveness of source-monitoring training more generalizable because the conditions of source monitoring are more consistent with everyday source judgments. Therefore, an important next step for research in both of these areas is to examine source-monitoring judgments about repeated events. Source monitoring is required in order to recall details specific to one event without confusing it with details from other events.

Children's memories for repeated events are qualitatively different from their memories for single events (see Roberts & Powell, 2001, for a review). When experiences are repeated, people generate a script for what usually happens, and this script is used to organize and make sense of common experiences (Nelson, 1986). For example, one might have a script for what happens at a restaurant (i.e., wait to be seated, order food, eat the food, etc.). The generic script would specify general aspects of the event and the order in which they occur, but would have different "slots" where one could fill in details specific to a particular event. For example, ordering food would be a general component of eating at a restaurant, but that could happen in different ways during specific events (e.g., ordering at a counter versus ordering from a waiter), so one of those choices could be "slotted in" to a specific event memory. Scripts are built up over time with repeated experience, but older children and adults can extract the key details of events to generate a script with less experience than can younger children. That is, young children take longer to create a script, and early on in script development, they are less aware of which details always belong to the script versus which details are deviations from the script (Farrar & Goodman, 1992; Brubacher, Glisic, Roberts & Powell, 2011).

Once a script has been formed it can aid recall by making the general event representation easier to recall. However, the exact details specific to one occurrence become more difficult to recall when experiences are repeated. This concept aligns well with Fuzzy Trace Theory (Brainerd & Reyna, 1990) in that the generic script is the “gist” of the event, whereas the specific details in each slot are verbatim traces. Gist representations are more accessible in memory and require less effort to retrieve, so as the script is rehearsed the gist of the event becomes better remembered, but the verbatim information may be lost. Compared with single events then, recognition decisions about repeated events may be easier because the script is reinforced through multiple experiences and identifying details from the general script only requires accessing a gist representation. However, source decisions about which specific details occurred during one event in particular are more difficult because there are multiple sources that may be highly similar, and source-monitoring errors often happen through confusions between events (called internal intrusion errors). For this reason, source decisions about repeated events are difficult for adults and children alike.

Because memories of repeated events are characteristically different from memories of single events, it is important to examine source-monitoring processes when making source decisions about a series of repeated events. However, in the past research on source monitoring and executive function as well as on source-monitoring training, there are no studies examining source monitoring of repeated events. The closest researchers have come to studying source-monitoring training with repeated events was when Roberts and Powell (2006) examined the effect of source-monitoring instructions on children’s reports about repeated events, but they administered the source questions *after* children freely recalled the event to test whether children would retract false reports, so their study did not involve explicit source training. Therefore, in Study 2, I explored the role of executive function and the effectiveness of source-monitoring training with respect to difficult source decisions: when the target sources were a series of repeated events and the source-monitoring task was to report about one occurrence from the series.

Using the same age groups as in Study 1, children aged 4-5 or 7-8 participated in a series of repeated events and were later asked to report about details from the last event. Each of the events had the same structure but the details varied in each event (e.g., there were storybooks at each event, but it was a different story each time). Prior to the interview, half of the children were randomly assigned to receive source-monitoring training before the memory test and the other half of the children received recognition training only. Executive functioning was measured through questionnaire responses from parents. The hypotheses were as follows:

1. It was hypothesized that there would be age differences in executive function, recognition, and source monitoring, consistent with the results of Study 1 as well as the extensive previous literature showing that older children outperform younger children on these tasks.

2. Because there were more than two sources (i.e., events) and the sources were highly similar, it was expected that recognition would be easy, but source decisions would be difficult; therefore, the role of executive function might be more evident for source monitoring than for recognition accuracy.

3. It was also expected that children in the source-monitoring training condition would demonstrate higher accuracy on the source-monitoring test than children in the control condition.

However, based on the results of Study 1, it was also hypothesized that an age by training condition interaction would reveal stronger source training effects for older children.

Method

Design

This study had a 2 (Age group: 4-5, 7-8) x 2 (Interview condition: Control, Training) between-subjects factorial design.

Participants

Initially, 131 children that were either 4- to 5 or 7- to 8-years-old were recruited for the study. In total, 35 children were excluded from the study; 27 missed one of the events, 5 were absent on the day of the interview, 2 were excluded due to interviewer errors, and 1 child exhibited a response bias (i.e., said yes to all details during the memory test, including those that never happened). This left a final sample of 96 children (51% male) with a mean age of 6.54 years (*SD* = 1.59). See Table 10 for sample sizes in each cell.

The socioeconomic status of participants was estimated by using the highest level of education obtained by the child's mother as a proxy. Generally the SES was high, with 82% of mothers attending post-secondary education; 42% had attended college, 34% had a Bachelor's degree, and 6% had a graduate degree. A high school diploma was the highest level of education for 13% of the sample, and 5% chose not to report education levels. Parents reported their children's ethnic background in response to an open-ended prompt, and their responses were later categorized. Ethnicity was not reported for 8% of the sample. Overall, 35% self-identified as Caucasian, 33% identified simply as "Canadian", 15% were Asian, 7% reported mixed ethnicity, and 2% were African American. The sample characteristics were similar to those of the sample in Study 1, as the studies were both conducted at elementary schools in the same region. However, none of the children in Study 2 had previously participated in Study 1.

Children were recruited from elementary schools in the local school board. Parental consent was obtained prior to children's participation in the study, and children also gave verbal assent at each session. Participants were compensated through a $10 donation to the school for each consent form returned.

Materials and Procedure

Children participated in five sessions; the first four were a series of similar repeated events, and the final session involved a training phase and memory interview about the last event. Parents also completed a questionnaire about their child's executive functioning.

Events. Over a two-week period, children participated in four sessions of a scripted 20 minute event that was referred to as the "Laurier Activities." Similar activities have been used in many previous repeated-event studies (e.g., Pearse, Powell & Thomson, 2003; Powell & Thomson, 2003; Roberts & Powell, 2005a; 2006). The same research assistant led each of the events, and a second research assistant helped to conduct the events and keep children focussed. Children participated in groups of up to 11 children. The average group size at the first event was 9.36 (*SD* = 1.69) and the average group size by the last event was 6.71 (*SD* = 2.43).

Each event was made up of 15 target items in the context of typical children's activities such as reading a story or doing a puzzle (Appendix C contains the script for an event). The structure and sequence of the events was the same each time, but the specific details varied across occurrences (see Appendix D for a list of instantiations). For example, children completed a puzzle of a clown at each event, but the puzzles were different each time: a clown painting, juggling, holding balloons, and standing under an umbrella across the four sessions. Four counterbalanced event orders were created so that each set of instantiations appeared equally as often in the first, second, third and fourth events. On the fourth occurrence, (the target event) all children wore a salient feather necklace that interviewers could use as a label to prompt the children at the interview (i.e., “I want to talk to you about the time you wore the feather necklace at the Laurier Activities”). This was the only occasion when children wore a necklace during the activities.

Training Phase. Between five and seven days after the fourth (and final) event, children individually participated in a training phase before completing a memory interview about the Laurier Activities. Each interviewer conducted interviews with children in both age groups and both interview conditions. For the training, children experienced two scripted activities (the Red Activity and the Blue Activity) and each of the activities had the same eight target details with two different instantiations across the events (e.g., wearing a red lei during the red activity and a blue lei during the blue activity). After the two activities the children in the control group completed recognition training and the children in the experimental group completed source-monitoring training.

In the control group, children were asked up to 12 recognition questions about whether details had happened during the Red and Blue Activities (the eight target details and four details not presented at the activities). They were given feedback about their responses (e.g., “You’re right, you did get a cat sticker!”). The interviewer never mentioned the sources in this condition. The children were asked recognition questions until they answered four consecutive questions correctly and then the interviewer proceeded to the target interview.

In the training condition, children were asked recognition and source questions (e.g., “Did you get the cat sticker during the Red Activity or the Blue Activity?”). Children were given feedback about their responses (e.g., “Actually, you got a cat sticker in the Blue Activity”). Children proceeded until they answered four recognition and source question pairs correct in a row, and then went on to the target interview.

Target Interview. Children were interviewed about the last time they participated in the Laurier Activities. The last time was chosen because recency effects should ensure that children remembered this event the best. The interviewer transitioned to the target interview by providing the following instructions: “Now it’s time to talk about the Laurier Activities. I wasn’t there when you did the Laurier Activities, and I need to know what happened the last time you did the Laurier Activities when you wore the feather necklace, so I’m going to ask you some questions about that time.” The interviewer confirmed with the children that they remembered that event. Children were asked 20 question pairs in random order about the last occurrence – 15 questions about the target details and 5 questions about details that never happened at the Laurier Activities. For each detail children were asked a recognition question first (e.g., “Did you do a puzzle the time with the feather necklace?”). If the children responded “Yes” to the recognition question, the interviewer followed up to determine which instantiation the children recalled for that occurrence (e.g., “What was the puzzle about the time with the feather necklace?”). Interviewers provided consistent generic positive feedback during the target interview (e.g., “You’re doing a great job!”) but did not provide feedback specific to the accuracy of responses. Because the events were fully counterbalanced, interviewers were not aware of which instantiations occurred during the last event for individual children.

Executive Function Assessment. Parents completed a paper copy of the BRIEF (Behaviour Rating Inventory of Executive Function; Gioia, Isquith, Guy & Kenworth, 2000; 2012) and returned it to their child’s school. Parents rated their children’s everyday behaviours over the last 6 months to indicate whether they displayed symptoms of poor executive functioning. This questionnaire contained 72 statements and each item was rated as occurring ‘never’, ‘sometimes’ or ‘often’. Subscales included Inhibition, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor. Appendix E contains a sample item from each subscale. Refer to Table 12 for the number of items in each subscale.

The validity of the BRIEF parent form in a normative sample has previously been demonstrated through a factor analysis of 1,419 participants’ scores, revealing two factors that accounted for 74% of the variance in the sample (Gioia et al., 2000). Emotional Control, Shift, and Inhibit scores loaded on one factor and the remaining five scores loaded on the other factor. Based on these findings, the BRIEF is used to calculate two composite scores, the Behavioural Regulation Index (BRI) and the Metacognition Index (MI). The questionnaire has been validated with children aged 5 to 18, but it was deemed appropriate for the 4-year-olds in this sample as well because all of the children had completed at least four months of formal schooling.

A parent questionnaire was chosen in part because the methodology was already very intensive with children participating in five sessions. Including a battery of cognitive tests would likely add two extra sessions for each child, and we anticipated problems with recruiting schools to participate in a seven-session study as well as problems retaining participants through to the end of the procedure. Additionally, using a parent questionnaire could help to resolve the problem of finding age-appropriate behavioural measures of executive function suited to a wide age range.

Coding

The questions in the target interview were scored for accuracy, and proportion scores were calculated for non-misleading recognition questions (correct identification of details that were present at the Laurier Activities) and false alarms (incorrect identification of details that were not part of the activities). Using the same approach as in Study 1, recognition accuracy scores were then calculated by subtracting the proportion of false alarms from the proportion of non-misleading recognition questions for each participant. “Don’t know” responses were conservatively coded as incorrect.

Source accuracy proportions were calculated by dividing the number of correct instantiations attributed to the last event by the total number of source questions asked (i.e., the number of correct identifications of details present at the activities). When children made source errors these were categorized as external intrusions (details that never occurred during the Laurier Activities), internal intrusions (instantiations that had occurred at the Laurier Activities during an event *other than* the last time) or “don’t know” responses. Proportion scores were calculated for each type of error. Internal intrusions were further categorized by which event the instantiations had come from, and proportion scores were calculated for the number of internal intrusions from the first event, second event and third event.

Each item on the BRIEF was numerically scored with a 1 if parents indicated “never a problem”, a 2 for “sometimes a problem”, and a 3 for “often a problem”. Therefore, higher scores indicated higher levels of executive dysfunction. Scores were summed on each subscale, and two composite scores (the BRI and MI) were also calculated for each participant by summing the scores on the relevant subscales.

The nature of the coding was objective (e.g., children’s yes/no responses were either correct or incorrect based on the details at the activities), so interrater reliability was greater than 99%. All of the data were coded by two raters to ensure accuracy. The few disagreements were due to addition errors and were resolved before data analysis.

Results

Analytic Strategy

Preliminary analyses were first conducted to identify outliers on the BRIEF scores, and assess the reliability and validity of the BRIEF. Preliminary analyses of the memory scores (recognition and source) tested whether the event leader, interviewer, event order, interview version, delay or gender were related to memory. Age, delay and executive function were also tested to verify that there were no differences between children in the two interview conditions. Next, the BRIEF scores were analyzed to assess developmental differences, as well as to determine whether executive function scores were related to recognition and source memory. Analyses were conducted to explore age differences in recognition and source monitoring, as well as differences between the control group and the children who received source-monitoring training. Finally, an exploratory analysis examined age and condition differences in the types of source-monitoring errors that children made. An alpha level of .05 was used to determine significance for all analyses.

Preliminary Analyses

BRIEF Scores. ***Outliers.*** Outliers on any of the eight subscales of the BRIEF were identified[2], and their composite scores involving those subscales were removed from further analyses (e.g., an outlier on the Inhibit subscale did not have a BRI score, but did have an MI score). It was deemed appropriate to remove the outliers because of concerns that they indicated either a response bias in parental reporting or unusually high levels of executive dysfunction; the present study examined executive function and source monitoring in a normative, not a clinical, sample. Seven children were outliers on one or more of the subscales, as shown in Table 11. Five

[2] Outliers were defined as any score more than 1.5 interquartile ranges below the first quartile or above the third quartile.

of these children were male and two were female. Five of these children were in the younger age group, and two were in the older age group.

Reliability. Cronbach's alpha was used to assess the reliability of each subscale, as well as the composite scores that were generated for each participant (BRI and MI). Reliability was good (above .70) for all individual subscales and the composite scores had very good reliability (see Table 12 for statistics on each subscale). Due to concerns about the applicability of the measure for younger children, reliability estimates were additionally calculated for the younger and older children's scores separately on each subscale. For most subscales the reliability estimates were very similar for both subsamples and in cases where the reliability estimates differed, the older children's scores consistently had lower reliability estimates. In most cases, even when the reliability estimates for the younger children were higher, the reliability estimates for both samples were greater than .70. However, the Initiate subscale showed reliability estimates of .77 for younger children and .65 for older children. The item-total statistics were examined and removing items 10 and 66 from this subscale improved reliability estimates to .72 for the older children, without substantially changing the reliability for the younger children (.73). Therefore, a new score was computed without those two items and composite scores involving that subscale were recalculated.

Validity of Composite Scores. To assess the validity of the BRI and MI scores in the present sample, a principal components factor analysis with a varimax rotation was conducted on the 72 items from the BRIEF. The results were very similar to those obtained in previous validation studies (Gioia et al., 2000). The analysis yielded two factors with eigenvalues greater than one, which accounted for 68% of the variance. The factor loadings confirmed that the BRI and MI composite scores were appropriate for the data given that Inhibit, Shift and Emotional

Control loaded on one factor, and the remaining five subscales loaded on the other factor. However, two subscales (Shift and Monitor) did load highly on both factors. Because the factor loadings for the Shift subscale were approximately equal for each factor, the scores on this subscale were considered problematic and were removed from further analyses involving the BRI. Although the Monitor subscale loaded heavily on both factors, the factor loading was substantially higher on Factor 1, as would be predicted by theory and previous validation studies, and therefore this subscale was included in further analyses of the MI[3]. Refer to Table 13 for the factor loadings. These two composite scores were used in all additional analyses involving executive function.

Face Validity. The face validity of some items on the BRIEF was questionable for the young age of the participants, and two parents indicated that they did not think some of the items applied to their child or that the questions seemed too mature for their child's age. Two raters that had extensive experience with young children examined the questionnaire and identified any questions that may not apply to the younger age group (4- to 5-year-olds). They agreed on eight items, some of which were those mentioned by the parents. Six of these items belonged to the 12-item Plan/Organize subscale (e.g., does not plan ahead for school assignments). One item was from the Monitor subscale, and one was from the Initiate subscale. Thus, it seemed that although the other subscales had acceptable face validity, the Plan/Organize subscale may not be valid for use with younger children. Because the Plan/Organize subscale contributed to the MI composite score, all analyses including the MI were re-run using an adjusted MI score that did not include the Plan/Organize subscale; the results of all analyses were the same, so analyses including the Plan/Organize subscale are reported.

---

[3] Note: All analyses were run with composite scores including the Shift and Monitor subscales, and then excluding these subscales. The results of all analyses were the same.

Distributions. The distribution of each composite score was examined to identify any issues with skewed data or a lack of variability in scores. Both the BRI and MI scores had means similar to the medians, indicating that the data were not skewed, and visual inspection of the histograms showed relatively normal distributions. See Table 14 for descriptive statistics for each of the subscales and composite scores. Refer to Figures 8 and 9 for histograms of the composite scores.

Recognition and Source-Monitoring Scores. Recognition accuracy was very good for most children, ranging from .32 to 1.00 with a mean of .87 (SD = .14). Source accuracy proportion scores were generally low, ranging from 0 to 0.75, with a mean of .33 (SD = .17).

Six counterbalancing measures were tested to ensure that they were not systematically related to either recognition or source memory. A series of one-way analyses of variance (ANOVAs) were conducted with recognition and source memory as the dependent variables. Independent variables included the event leader (one of four female research assistants), interviewer (one of four female research assistants), event order (one of four counterbalanced orders as described in the method section), and interview version (one of three interviews with similar activities). All were non-significant, Fs ≤ 2.51, ps $\geq .06$, $\eta^2_p \leq .08$. Independent-samples t-tests confirmed no effect of the order of the two training events on recognition or source memory, ts ≤ 1.86, ps $\geq .07$, Cohen's ds ≤ 0.38. Two one-way (Delay: 5, 6 or 7 days) ANOVAs revealed no significant difference in either recognition or source memory depending on the delay between the last event and interview, Fs ≤ 2.32, ps $\geq .10$, $\eta^2_p \leq .05$. There were no gender differences in either recognition or source memory, as assessed by independent-samples t-tests, ts $\leq |1.30|$, ps $\geq .20$, Cohen's ds ≤ 0.27. Any variables that were marginally significant were tested as covariates in further analyses, but none were significant. Thus, analyses are reported without covariates.

Although children were randomly assigned to interview conditions and personal characteristics should statistically even out across conditions, preliminary analyses were conducted to ensure that there were no differences between the children assigned to the control and training conditions in age, delay between the last event and interview, or executive function scores. These variables were expected to influence memory; thus, any differences between children in the control and training conditions would need to be taken into account when drawing conclusions about the effects of interview condition. A series of independent-samples *t*-tests confirmed no differences between the control and training conditions on average age, delay, or any of the 10 executive function subscale or composite scores, $t\text{s} \leq |1.23|$, $p\text{s} \geq .22$, Cohen's $d\text{s} \leq 0.26$. An additional independent-samples *t*-test confirmed no differences between the younger and older age groups in the average delay, $t\ (94) = -0.11$, $p = .91$, Cohen's $d = 0.02$.

Training. Of the 96 children in the sample, 87 passed the criterion of four consecutive correct questions/question pairs before proceeding to the target interview. Nine participants were asked all 12 questions/question pairs but never met the criterion before proceeding to the target interview. All but one of those children were in the younger age group, and seven were in the source-monitoring training group, whereas the other two were in the control group. For those that did meet the criterion, it took five questions on average to get four correct in a row ($M = 4.97$, $SD = 1.76$). As in Study 1, all participants were included in further analyses, regardless of whether they passed the criterion or not.

Inferential Analyses

Developmental Differences in Executive Function. The BRI and MI scores were significantly correlated, r (67) = .53, p <.001. Therefore, in order to assess developmental differences in executive function, a one-way (Age Group: 4-5, 7-8) multiple analysis of variance (MANOVA) was conducted on the BRI and MI scores. There was a significant multivariate effect of age on the scores, F (2, 66) = 9.25, p < .001, Wilk's $\lambda = 0.78$, $\eta^2_p = 0.22$. Two follow-up independent samples t-tests for the BRI and MI scores separately showed a significant age difference for BRI scores, t (53.44) = -2.82, p = .007, Cohen's d = 0.54, but not for the MI scores, t (70) = -.40, p = .69, Cohen's d = 0.10. Younger children had significantly higher BRI scores (indicating poorer behavioural regulation) than did older children ($M_{\text{younger}} = 34.10$, SD = 6.76; $M_{\text{older}} = 30.00$, SD = 4.96). To confirm age differences this analysis was also run using two linear regression with age as the predictor and BRI and MI scores as the dependent variables. Age was not a significant predictor of the MI scores, F (1, 72) = .62 , p = .44, but was a significant predictor of BRI scores (F (1, 68) = 7.13, p = .009).

To examine the age differences in the BRI more closely, a one-way (Age Group: 4-5, 7-8) MANOVA was conducted on the subscale scores that contributed to the composite BRI scores (Inhibitand Emotional Control). Again, there was a significant effect of age on the scores as a group, F (2, 67) = 4.30, p = .02, Wilk's $\lambda = 0.89$, $\eta^2_p = 0.11$. Independent samples t-tests were used to examine age differences in each dependent variable separately. These tests revealed significant age differences on both the Inhibit and Emotional Control subscales, ts ≥ 2.54, ps $\leq$ .01, Cohen's ds ≥ 0.61. See Table 15 for the means on each of the three subscales compared across age groups.

The Relationship between Executive Function and Memory. It was expected that scores from the BRIEF would be related to both recognition and source-monitoring performance. Regression analyses were run separately for the two dependent variables. Both the MI and BRI scores were entered simultaneously to assess whether they predicted outcome variables independent of one another.

For recognition memory the model was not significant, F (2, 66) = 1.34, p = .27, R^2 =.04. Neither the BRI nor MI scores were related to how well children could identify the items that had occurred during the last event.

For source accuracy the model was significant, F (2, 66) = 3.78, p = .03, and accounted for 11% of the variance in source-monitoring scores. Both the BRI and MI scores were independently relatedto source-monitoring accuracy. Executive function was related to how well children could identify the instantiations from the last event without confusing instantiations from the previous three events. When age was added as an additional independent variable, neither the BRI nor MI scores remained significant. The overall model was significant, F (3, 65) = 8.90, p < .001, R^2 = .29, but age was the only variable to make a significant contribution. Standardized regression coefficients and their associated test statistics can be found in Table 16.

Age and Interview Condition Differences in Memory. It was expected that older children would have higher recognition and source accuracy than younger children. It was also expected that children in the source-monitoring training group would outperform those in the control group with respect to source-monitoring accuracy. Because recognition and source accuracy scores were significantly correlated, r (94) = .44, p <.001, a 2 (Age Group: 4-5, 7-8) x 2 (Interview Condition: Control, Training) MANOVA was used to assess the effects of age group and interview condition on both dependent variables. There were significant effects of Age

Group, F (2, 91) = 20.77, p < .001, Wilk's $\lambda = 0.69$, $\eta^2_p = 0.31$, and Interview Condition, F (2, 91) = 3.53, p = .03, Wilk's $\lambda = 0.93$, $\eta^2_p = 0.07$, on the dependent variables as a group.

Follow-up 2 (Age Group: 4-5, 7-8) x 2 (Interview Condition: Control, Training) ANOVAs for each dependent variable clarified the multivariate effects. For recognition scores, there was a main effect of Age Group, F (1, 92) = 22.55, p < .001, $\eta^2_p = 0.20$, a marginal main effect of Interview Condition, F (1, 92) = 3.44, p = .07, $\eta^2_p = 0.04$, and a marginally significant interaction, F (1, 92) = 2.63, p = .10, $\eta^2_p = 0.03$. Older children (*M* = .93, *SD* = .09) had higher recognition accuracy scores than younger children (*M* = .80, *SD* = .16). The marginal effect of condition revealed that children in the training condition (*M* = .89, *SD* = .10) had slightly higher recognition accuracy than children in the control group (*M* = .84, *SD* = .18). Examination of the means demonstrated that the difference between the control and training groups was much larger for the younger children, explaining the trend towards a marginally significant interaction effect. See Table 17 for mean recognition scores by Age Group and Interview Condition.

For source accuracy scores, there was a main effect of Age Group, F (1, 92) = 31.00, p < .001, $\eta^2_p = 0.25$, but no effect of condition or interaction, *F*s < 2.49, *p*s > .12, $\eta^2_p \leq 0.03$. Again, older children were more accurate at recalling instantiations from the last event (*M* = .41, *SD* = .15) than younger children (*M* = .24, *SD* = .15). One-sample *t*-tests compared the younger and older children's mean source accuracy scores to chance probability, which was set at 0.25 (i.e., a 25% chance of answering correctly because there were four instantiations of each detail). The older children were performing significantly above chance on this source-monitoring task, t (49) = 7.39, p < .001, but the younger children's scores did not differ from chance, t (45) = -.43, p = .67.

Given the high rate of errors, an exploratory analysis examined the types of errors that children made across age groups and interview conditions. A 2 (Age Group: 4-5, 7-8) x 2 (Interview Condition: Control, Training) x 3 (Error Type: Internal intrusions, External intrusions, "Don't Know") mixed ANOVA with repeated measures on the last factor revealed significant effects of Error Type, F (1.46, 134.17) = 73.08, $p < .001$, $\eta^2_p = 0.44$, an Age Group by Error Type interaction, F (1.46, 134.17) = 10.02, $p < .001$, $\eta^2_p = 0.10$, and a significant three-way interaction, F (1.46, 134.17) = 3.94, $p = .03$, $\eta^2_p = 0.04$. See Table 18 for the mean proportions of each error type by Age Group and Interview Condition.

To better understand the three-way interaction, follow-up 2 (Interview Condition: Control, Training) x 3 (Error Type: Internal intrusions, External intrusions, "Don't know") ANOVAs with repeated measures on the last factor were run for younger and older children separately. For younger children there was only a main effect of Error Type, F (1.44, 63.27) = 11.32, $p < .001$, $\eta^2_p = 0.21$, and no interaction with Interview Condition, F (1.44, 63.27) = 1.51, $p = .23$, $\eta^2_p = 0.03$. Younger children made mostly internal intrusion errors, followed by "don't know" responses, and external intrusions were least likely. Interview condition had no effect on the types of errors that younger children made.

For older children, there was a significant effect of Error Type, F (1.37, 65.52) = 101.39, $p < .001$, $\eta^2_p = 0.68$, as well as an Error Type x Interview Condition interaction, F (1.37, 65.52) = 3.10, $p = .05$, $\eta^2_p = 0.06$. Compared to the control group, older children who had received the training were more likely to respond with "don't know" instead of mistakenly reporting an instantiation from an event other than the last one.

Finally, for those children who made internal intrusion errors, an additional analysis examined which event in the series the instantiations came from and assessed whether there were age or training effects on this aspect of children's responding. A 2 (Age Group: 4-5, 7-8) x 2 (Interview Condition: Control, Training) x 3 (Source of Internal Intrusions: First event, Second event, Third event) mixed ANOVA with repeated measures on the last factor was conducted. There were main effects of the event that intrusions came from, F (1.65, 147.40) = 49.78, p < .001, η^2_p = 0.36, and a marginal interaction with age, F (1.65, 147.40) = 3.04, p = .06, η^2_p = 0.03, but there was no interaction with interview condition, F (1.65, 147.40) = 0.92, p = .39, η^2_p = 0.01. Post-hoc *t*-tests with a Bonferonni correction showed that overall children reported the most internal intrusions from the third event (*M* = .56, *SD* = .26), followed by the second event (*M* = .28, *SD* = .24), and then the first event (*M* = .16, *SD* = .17), indicating that temporal proximity to the to-be-remembered event played a role in source confusions. The marginal interaction with age showed that the pattern was the same for the younger and older age groups, but the older children reported fewer instantiations from the first and second events and more instantiations from the third event, relative to the younger age group.

Discussion

The present study used a repeated-event paradigm to study the relationship between executive function and source monitoring, as well as to examine the effect of source-monitoring training among 4- to 8-year-old children. Participants experienced a series of four similar scripted events consisting of 15 target items, which had different instantiations each time. Children were randomly assigned to interview conditions so that half of the children received recognition questions and the other half received source-monitoring training, before completing a source-monitoring test in which they were asked to recall the instantiations from the last event. Parents also completed the BRIEF to measure children's executive functioning. This is the first study to examine the executive function-source monitoring relationship or source-monitoring training in the context of source decisions about repeated events.

Developmental Differences

Developmental differences were expected on the BRIEF, as well as for the recognition and source-monitoring scores. The results showed age differences on the Behavioural Regulation Index (BRI) of the BRIEF, but not the Metacognition Index (MI). Specifically, the Emotional Control and Inhibition subscale scores were both lower for older than for younger children, indicating that older children had better emotional control and inhibition. Because age differences were expected for both components of executive function, it was surprising that the younger children scored just as well as the older children on the MI. In particular, it was expected that individual differences in working memory would be related to age based on the results of Study 1. However, working memory was only one of five subscales contributing to the MI scores, so perhaps some of the other subscales did not show age differences, and therefore overall the MI was not significant. The other subscales may not have been related to age because these skills develop in later childhood, and our sample was too young to detect developmental differences. Another possible explanation for the difference in findings between studies is that behavioural measurements of working memory may be more sensitive than parent ratings, and therefore, they relate more closely with age.

It is also worth noting that all of the questions that posed challenges to the face validity of the questionnaire for the younger age group were from the MI component of the BRIEF. The subscales contributing to the MI are Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor. When parents of younger children found that an item did not apply to their child, they may have indicated that these items were “never a problem”, which would serve to limit the variability of the sample and minimize age differences. Regardless, the MI scores showed significant relationships with other variables of interest (see below), so meaningful analyses of the MI composite scores were achieved, even though they did not vary with age.

Age differences were present in both recognition and source accuracy. The older children were better than the younger children at recognizing which items had occurred during the last event, as well as recalling instantiations specific to the last event. The relatively high recognition scores in both age groups (proportion scores of .80 for younger children and .93 for older children, on average) demonstrate strong script development. When children experience events that are highly similar across occurrences, they develop scripts for what usually happens (Farrar & Goodman, 1992). As scripts become stronger, their representations of individual events become more script-like, and it becomes easier to identify which items were present at the events by comparing to the script or using the script to facilitate recall (Pearse, Powell & Thomson, 2003). Because children experienced four similar events with the same items each time and varying instantiations of each item, they developed a script for which items occurred during the Laurier Activities. Therefore, their recognition for the items that were or were not present at the activities was good. The age differences in recognition demonstrate that the younger children's scripts were not as well developed as the older children's after only four occurrences of an event, which is consistent with literature showing that younger children need more experience with a repeated event in order to generate a script (e.g., Farrar & Goodman, 1992).

In terms of source accuracy, mean scores for both age groups were quite low. The younger children were very poor at attributing the instantiations from the last event and performed at chance levels. The older age group performed better than chance, and significantly better than the younger age group, but still, the mean proportion accuracy score for the older group was only .41. Again, strong script development can help to explain the children's performance; well-developed scripts tend to promote more generic, schema-based recall with fewer details about individual occurrences. As experiences are repeated, it becomes more difficult to distinguish which time unique instantiations occurred (Pearse, Powell & Thomson, 2003). These results demonstrate the difficulty of source-monitoring judgments about repeated events that are highly similar, highly scripted, and temporally close together. In line with the large literature on memory for repeated events more broadly (e.g., Powell & Thomson, 1996; Powell & Thomson, 2003), internal intrusion errors (i.e., confusions between events) were common, and were the main sources of error for both younger and older children (as opposed to external intrusions or "don't know" responses).

The findings regarding good recognition accuracy and relatively poor source-monitoring accuracy also align well with Fuzzy-Trace Theory. The recognition questions required knowledge of the items that were present during the activities, or what would be considered the gist of the events, whereas the source questions required retrieval of specific instantiations from one event, encoded at the verbatim level. Because gist traces are more accessible in memory and require less effort to retrieve (Brainerd & Reyna, 1990), from the perspective of Fuzzy-Trace Theory, it is not surprising that recognition scores were much higher on average than source-monitoring scores.

Although Script Theory and Fuzzy-Trace Theory provide compelling explanations for why the recognition scores were much higher than the source-monitoring scores, a third explanation hinges on the structure of the questioning for the two tasks. The recognition test consisted of yes-no questions which required children to identify items that were present at the Laurier Activities or not (e.g., "Was there a story?"), whereas the source-monitoring test consisted of cued recall questions that required children to generate a response (e.g., “What was the story about?”). Decades of memory research, including studies using repeated-event paradigms, have consistently shown that recognition tasks are easier than recall tasks (e.g., Craik & McDowd, 1987; Gillund & Shiffrin, 1984) so this could be an alternate explanation for these results. A source-monitoring test using forced-choice questions (e.g., “Was the story about a birthday party or a boat?”) would have allowed for a more carefully controlled comparison of recognition and source accuracy, but would also have made the procedure less generalizable because in real-world interview scenarios, a naïve interviewer would not be able to offer a list of instantiations to choose from. Because a comparison of recognition and source memory was not a main goal of the present study, we opted for a more realistic questioning style.

Regardless of whether the reason for poor source-monitoring performance is script-based recall, loss of verbatim traces, or difficulty responding to cued recall questions, one thing is clear: these findings demonstrate the difficulty of these types of source judgments as they occur in everyday remembering, and highlight the importance of research in pursuit of interview techniques that help children overcome these difficulties.

When children made internal intrusion errors by reporting instantiations from another event, most intrusions came from the third event, followed by the second event, and finally the first event, which demonstrated that temporal proximity to the to-be-remembered event is a factor in decision-making. A marginal interaction with age showed that the older children tended to intrude instantiations that were closer to the target event in terms of temporal order (e.g., the third event rather than the first), indicating that not only was their accuracy better for the last event compared to younger children, but when they did make mistakes, they were closer to the correct instantiations in temporal order.

These results provide some interesting insight into the cognitive processes behind source-monitoring decisions. One explanation is that when it comes to repeated events, these decisions rely heavily on the temporal order and sequencing of the events. Children may attempt to recall the order of events to assist in their decision-making, and be more likely to confuse the third and fourth events than the first and fourth events because the latter are farther in order. Young children struggle with providing temporal information about events (Friedman, 1993), and this could be one reason why the older children were more successful than the younger children at this difficult source-monitoring task.

Another explanation could be that children rely on cues about the characteristics of their memories, such as the amount of perceptual detail, in order to judge how long ago an event happened (Friedman, 1993; Friedman, 2004). The third and fourth events would be more similar with respect to such characteristics than would the first and fourth events, which would explain the tendency to intrude details from events that were temporally closer to the last event. This would support Johnson et al.'s (1993) proposition that source decisions can be made by comparing the characteristics of memories from different sources, and would therefore lend support to the Source-Monitoring Framework. From this perspective, age differences would occur because older children have better knowledge of strategies such as comparing source characteristics; better developed cognitive skills, such as executive function, to assist in carrying out such strategies; or more accurate metamemory assessments.

Executive Function and Memory

It was expected that the role of executive function would be more evident with respect to source-monitoring accuracy than it would for recognition accuracy. The BRI and MI scores were not significant predictors of recognition accuracy, perhaps because, as described above, the recognition task was much easier than the source-monitoring task. Executive function may be less implicated in memory decisions that are less cognitively demanding (such as recognition judgments) because these decisions can be made quickly and easily, and complex cognitive strategies such as those required for difficult source decisions are not necessary (Johnson et al., 1993).

Both the BRI and MI were significant independent predictors of source accuracy. This demonstrates the role of various components of executive function in source-monitoring processes, as the BRIEF measures not only inhibitory control and working memory, but also other aspects such as monitoring, planning, and organization. There is still debate in the field more broadly about the structure and components of executive function, but this study extends the evidence to show relationships between source monitoring and executive function beyond working memory, inhibitory control, and cognitive shifting, which have been the focus of previous literature. Also, the vast majority of previous studies relating executive function to source monitoring have used behavioural measures, so finding significant relationships using a parent-report is encouraging. This demonstrates that the relationship goes beyond a specific methodology and also provides some support for the validity of the BRIEF as a measure of executive function.

One caveat was that neither the MI or BRI scores remained significant predictors of source accuracy when age was entered into the regression. This may indicate that a global developmental factor could explain source monitoring more fully than executive function (as measured by the BRIEF) alone. On one hand, perhaps this should not be surprising, as there are many cognitive developments other than executive function that happen between ages 4 and 8. The global cognitive development factor represented by age includes improvements in frontal lobe development that affect executive function as well as many other domains, and therefore, it should be a stronger predictor of source monitoring. On the other hand, this is contradictory to the results of Study 1, which showed that executive function predicted source monitoring even when controlling for age, and this is explored further in the General Discussion.

Source-Monitoring Training Effects

It was expected that the children in the source-monitoring training condition would outperform the children in the control condition with respect to source accuracy for recalling instantiations from the final occurrence of the event, and that the training effects may be larger for older children. Surprisingly, there were no effects of training on source accuracy for either age group. The children who received the source training were no more accurate with respect to identifying specific instantiations than those who were in the control group. Although older children performed better than chance on the source-monitoring test and therefore had the best chance to improve their source monitoring relative to the younger age group, their scores were still low on average and training did not make a difference in increasing accurate source monitoring as measured through cued recall questions.

However, source training did have an impact on the types of errors that the older (but not younger) children made. As discussed above, the most common source errors for both age groups were internal intrusions, followed by “don’t know” responses, and lastly external intrusions. Within the older age group, children who received source training were more likely to err by saying “I don’t know” and less likely to make internal intrusion errors than the children in the control group. In other words, they showed a heightened awareness of the comparison between events and less tendency to confuse instantiations. Although the training did not improve the number of accurate responses, this finding seems to indicate that it made children more conscious of the repeated nature of the events and the difficulty of the source decisions, and children in the training group metacognitively monitored their knowledge more accurately.

It is possible that these findings demonstrate that developmental changes are occurring, and that source-monitoring performance happens on a continuum. First children perform poorly on source-monitoring tasks because they lack the cognitive skills or strategies to source monitor effectively. This is followed by a stage of heightened metacognitive awareness of the requirements of the task, but a lack of confidence in responses. This stage may also be characterized by a utilization deficiency, in which children begin to use strategies, but their performance does not benefit. Finally, children progress to a point where they have the ability to use strategies effectively to improve their source monitoring, and perform at similar levels to adults.

If this is the case, the results of this study demonstrate that for the older children in the sample, the source-monitoring training did scaffold their source-monitoring performance from the first to the second level. The older children who received training were more likely to say “don’t know” than make internal intrusion errors, relative to the control group, demonstrating that these children were more aware of the requirements of the source-monitoring task and were monitoring their memory decisions more carefully. However, the benefits did not translate into more accurate source-monitoring performance because children were not yet ready to use these strategies independently. The training made these children more aware of what they were required to do, but not better able to do it. The older children’s don’t know responses could potentially represent a developmental pathway from inaccurate source-monitoring to accurate source-monitoring. Because of the difficulty of the source decisions in this particular study, the children were not able to improve their accuracy. They began to use strategies, but the strategies were not particularly useful to their performance.

The Source-Monitoring Framework would explain these findings as the children adjusting their criterion for making source decisions. Johnson et al. (1993) proposed that people set a criterion for making a judgment based on factors such as the importance of the decision or how confident one feels about the memory characteristics. Criteria can change based on situational factors, such as goals or motivation. Children's increased tendency to say "I don't know" after training shows a stricter criterion for what they would accept as belonging to the last event. Because the training highlights that source information is important to the interviewer, children potentially held back answers that they felt less confident about, which resulted in more "don't know" responses and fewer internal intrusions (i.e., source errors). From a practical perspective, the findings are encouraging. For high-stakes source decisions such as providing eyewitness testimony, stringent criteria for source decisions are desirable because an incorrect source attribution could be more damaging for both the investigation and the child's credibility than the potential loss of information associated with a "don't know" response.

The source decisions in the present study were very difficult, as evidenced by the low mean accuracy scores for both age groups and the fact that the younger group responded at chance levels. When source decisions are this difficult, training effects may be harder to find because the source-monitoring decisions are beyond young children's developmental capabilities and they are not able to improve, even with an intervention targeting strategy use. For example, as described above, it is possible that the children relied on temporal information about the sequence of events when making their source decisions. If younger children do not have the temporal knowledge to carry out this type of decision-making process, it is unlikely that training could improve their accuracy.

Although the difficulty of the source decisions potentially limited the results, it was necessary to study source-monitoring training under these conditions because the goal of the present research is to generalize to conditions such as forensic interviews, when children are often interviewed about repeated events after long delays. If the source training is not effective when source decisions are very difficult, it has limited real-world value because those are the conditions under which source decisions are made in everyday remembering.

Limitations and Future Directions

One limitation of the present research was that executive function was measured in different ways across studies, which limits comparisons of the results. However, significant relationships between executive function and source monitoring were found in both studies, regardless of the methodology. Future research could examine behavioural measures and parent reports of executive function within the same study in order to show the relative contributions when both are used together. This area of research requires further exploration in order to determine the best ways of measuring executive function accurately.

Assessing the training using a repeated-event paradigm with difficult source decisions allowed for a better evaluation of the practical usefulness of the interview technique compared to previous studies that have used easier source decisions involving perceptually distinct sources after short delays. Future research can continue to examine the effectiveness of source training under varying conditions of task difficulty. It is essential that these results be replicated before recommendations can be made to practitioners, and an important area for future study is to expand research on task difficulty, for example, by using more similar versus more different sources. This could provide more insight as to whether the effectiveness of the training changes under more difficult circumstances for source monitoring.

Further research could also examine the same training procedure with an older age group to determine whether there might be benefits of training for source accuracy. It is possible that when children are more developmentally ready for using strategies to source monitor in difficult circumstances, the training could scaffold children to source monitor more accurately than their peers in the control group.

General Discussion

Because the two studies in this dissertation addressed similar research questions, the results can be considered together in the greater context of the literature and theory in this field. Two key areas were considered: executive function, and source-monitoring training, with respect to difficult source-monitoring decisions.

Executive Function

In order to examine the relationship between executive function and source monitoring, this dissertation used a variety of methods, including easier and more difficult source decisions, source monitoring with two perceptually distinct sources versus a series of repeated events, and behavioural measures of executive function as well as parent reports. The finding that executive function was related to source monitoring across both studies provides strong support for its role in making accurate source-monitoring attributions because the relationship goes beyond a specific methodology.

Interestingly, in both studies, working memory and inhibitory control were identified as two separate factors that independently predicted source monitoring. In Study 1, a factor analysis confirmed that working memory and inhibitory control scores were independent, and in regression analyses both contributed to the variance in source monitoring. Although the BRIEF sub-component scores also measured other aspects of executive functioning, the BRI score included inhibition and the MI score included working memory, and again, these were found to be significantly related to source monitoring. Taken together then, these studies provide evidence not only that executive function plays a role in source monitoring, but also that individual components of executive function contribute separately. In addition, Study 1 supports the importance of the most widely agreed upon components of executive function: working memory and inhibitory control. The evidence from mediation analyses incorporating both executive function and encoding into a model of source-monitoring development is an important contribution to the current literature in the field because in addition to showing support for the role of executive function, it explains *how* executive function exerts its influence over children's developing source monitoring skills. Specifically, improvements in working memory lead to better event encoding, and subsequently more accurate source monitoring.

Despite these very encouraging findings, it is important to note that in both studies, regression models using executive function variables to predict source monitoring could only account for 11% of the variance in source-monitoring scores. Thus, although there is compelling evidence that executive function plays an important role in children's source-monitoring development, there are many other factors (including both cognitive and interview-related factors) that also play a role. This is an area requiring future research, as other individual difference variables and situational factors can be explored to determine how they work together to impact source monitoring.

Generally the results of the two studies complement one another, but there were also some key differences in the findings. First, in Study 1, executive function was related to recognition, whereas this was not the case in Study 2. As discussed earlier, it is likely that children's script development impacted the recognition scores in Study 2, as it would be quite easy for children to identify the gist of the events, separate from having to recall specific instantiations. This greatly reduced the cognitive demands of the task compared to the recognition task in Study 1, which comprised questions about specific details from a novel, unique event that was only presented once. Because the cognitive demands were so much higher in the more difficult recognition task in Study 1, this helps to explain why executive function may have played a stronger role in recognition judgments for Study 1 than Study 2. The majority of previous studies of executive function and source monitoring have used a combined recognition/source task to measure memory (e.g., Asking, "Was there a puzzle in the real-life demonstration?" instead of asking separate questions about the presence of a puzzle, and the source of the puzzle; Roberts & Powell, 2005b; Karpinski & Scullin, 2009). This dissertation allowed for a separate examination of the relationship of executive function to recognition and source monitoring, and although the results are mixed, overall there is evidence that in some circumstances, executive function can predict recognition accuracy as well.

A second difference between the findings of the two studies was that the relationship between executive function and source monitoring held true when controlling for age in Study 1, whereas this was not the case in Study 2. This difference in results might be attributable to differences in the way that executive function was measured between the two studies. Specifically, in Study 1 behavioural measures of executive function may have been more precise than the parent report used in Study 2, which asked parents to describe their children's executive functioning over the last 6 months. Using behavioural measures allowed us to pinpoint the children's executive functioning at an exact point in time, which was a stronger predictor than age, which represents development over the span of an entire year. On the other hand, the BRIEF captured a snapshot of the children's executive functioning over a 6-month period, and because so much can change in a child's development over a span of 6 months, it is not surprising that exact age would be a more accurate predictor than the BRIEF, even if there is shared variance.

In summary, this dissertation provides evidence for the role of executive function in children's source-monitoring development. However, there are many other research questions in this area that still need to be addressed. An important next step is to continue to examine interview-related factors to determine how they affect the role of executive function in source monitoring. In the present research, controlling task difficulty by manipulating the delay period did not lead to differentiated relationships between executive function and source monitoring, but manipulating task difficulty in other ways (e.g., through source similarity or the type of source-monitoring task [i.e., internal versus external source decisions]) may lead to different conclusions.

Source-Monitoring Training

One of the major goals of this dissertation was to examine the effectiveness of source-monitoring training in the context of task difficulty and repeated events. Although previous studies on source monitoring have used a variety of easier and more difficult source-monitoring tasks, no study to date has compared the effectiveness of the training with varying task difficulty. Similarly, aside from *post*-recall source-monitoring instructions (Roberts & Powell, 2006), source-monitoring training has not been examined when the source decisions require distinguishing between a series of repeated events.

To summarize the results, in Study 1, training effects were qualified by an interaction with age and delay; the only group to benefit from the training was older children in the shorter delay condition. In Study 2, there were no effects of training on source accuracy, but training did affect the types of source errors that older children made, such that they were less likely to make internal intrusion errors and more likely say “I don’t know”. One finding common to both studies was that source-monitoring training did not have any effects for the 4- to 5-year-old age groups.

It was surprising that source training did not impact the accuracy of source-monitoring decisions in Study 2 as it had in Study 1. Children’s source errors in Study 2 were most commonly internal intrusion errors, in which they confused instantiations from another event when trying to recall the last event in the series. The fact that children were able to generate instantiations from other events shows that it is not their memory for verbatim details that is the problem, as would be expected based on Fuzzy-Trace Theory, but rather the source attribution process, as predicted by the Source-Monitoring Framework. Children were able to generate many verbatim details from the events, they just could not make judgements about which had occurred in the last event.

One possible explanation for this is that the training in Study 2 involved only two sources; the training consisted of completing two non-target events and a source-monitoring test about these events with feedback. The training events were similar to the Laurier Activities in their structure and each event had the same items, but with different instantiations in each. However, it is possible that the training children completed with two sources would not transfer to source decisions involving more than two sources. Another possible explanation is that because the source decisions were so much more difficult in Study 2, the children could not improve their source monitoring even with the training. The task was so hard for the young children that their scores were no better than chance, and instructions that draw attention to the importance of source or highlight strategy use are not helpful. For older children there was evidence that the training caused them to consider source more carefully, hence their increased tendency to say "I don't know" instead of making internal intrusion errors. However, it did not improve their accuracy overall.

Although the training effects were clearly complex and context-dependent, both studies found that the younger age group (4- to 5-year-olds) did not benefit from training. This is consistent with the work of Poole and Lindsay (2002) who have found that older, but not younger, children benefit from training. This is in contrast to the work of Thierry and her colleagues (Thierry et al., 2010; Thierry & Spence, 2002; Thierry et al., 2001), who have found training effects for children as young as 3- to 4-years-old. One of the differences between previous studies is the delay period, which was the focus of the present research. Study 1 isolated that one variable, which reliably affects memory and source-monitoring, in an attempt to clarify if that factor was contributing to the inconsistent results of previous studies. The results of Study 1 were not as illuminating as expected, but there are many other differences between these studies that could be responsible and deserve further consideration.

One example of another difference between Poole and Lindsay (2002) and Thierry and Spence (2002) is that while Poole and Lindsay trained the children on non-target events, Thierry and Spence's source-monitoring training involved a free recall interview about the target events that they would later be asked source-monitoring questions about. In other words, Poole and Lindsay's training required that children learn a strategy and then transfer that skill or knowledge to a new scenario, whereas Thierry and Spence's training did not require that the strategy be used with regard to a separate event. Perhaps this could explain why younger children only benefit from training in Thierry et al.'s work; if children learn a new strategy they can use it within the same task, but are not able to transfer their learning to a new task. This is an important area for future research in terms of isolating factors that could potentially explain why the findings have been so inconsistent in past studies.

One important finding to note was that in both studies, 10-20% of the participants did not pass the criterion for training. Because the cell sizes were small for participants that did not pass the criterion, no meaningful analyses could be conducted comparing them to children who did pass the criterion, but it is unclear whether these children in particular would benefit from the training at all. An important area for future research would be to combine research on source training with research on individual difference factors (such as executive function) to predict which children would benefit from the training and why. It may be the case that certain cognitive skills are prerequisite in order to benefit from the training, and exploring these further would help to clarify the circumstances when source training would be helpful.

From a practical perspective, source training shows promise because the technique could be used by anyone – no special training or qualifications would be required. However, the practical implications for use in the field are limited at this point because of the mixed evidence for the usefulness of this interview procedure. We are a long way from fully understanding what conditions affect the efficacy of the training procedure, so it is essential that researchers continue to systematically explore the circumstances that affect whether children benefit from the training or not.

Conclusion

The goal of this dissertation was to examine cognitive and interview factors that affect young children’s source-monitoring abilities. The first main research question involved clarifying the relationship between executive function and source monitoring by testing the relationship with a variety of methods, including easier and more difficult source decisions, source monitoring two discrete sources versus a series of repeated events, and measuring executive function through behavioural measures and parent reports. Taken together, the findings from both studies show strong evidence that executive function does play a role in children's developing source-monitoring skills. Not only that, but the results also provide more insight into *how* source-monitoring skills develop across early childhood. Study 1 demonstrated that as children get older their working memory progresses, which leads to subsequent improvements in event encoding, and in turn, source monitoring. These results are the most conclusive evidence to date relating executive function to source monitoring, and especially to developmental improvements in source monitoring. Surprisingly, task difficulty did not influence the way that the strength of the relationship presented, and therefore is not likely to be a contributing factor in the contradictory results of previous work. Future research should continue to explore what external factors cause the relationship to present differently in an attempt to further explain the results of previous literature.

The second main research question involved identifying circumstances in which source-monitoring training is or is not an effective means of improving young children's source accuracy by examining it's effectiveness with easier and more difficult source decisions, and also within the context of repeated events. Across the two studies, training effects proved to be complex and qualified by other variables. Study 1 found evidence that older children's accuracy would improve with training in the shorter delay condition, and Study 2 found that older children who received training considered source more carefully and were more likely to respond "I don't know" instead of making source errors, but neither study found evidence of training effects with younger children. Future research can replicate these results by manipulating task difficulty in other ways, such as by altering the similarity of the sources involved in source decisions. In addition, continuing to isolate methodological differences between previous studies will provide more insight into both the mixed results in the literature, and more broadly, the factors that influence the effectiveness of the training.

Rigorous scientific testing of these relationships in a variety of conditions provided further information about how cognitive and interview factors are related to children's source monitoring. Collectively, my doctoral program of research contributes a greater understanding of how source monitoring develops and the circumstances when source-monitoring training is effective.

Table 1

Sample Sizes within Each Cell for Study 1

Age Group	Delay Condition	Control Condition	Training Condition
4-5 years	Shorter	23	24
	Longer	24	24
7-8 years	Shorter	24	24
	Longer	23	24

Table 2

Factor Loadings of Executive Function Measures on Two Factors

Measure	Factor 1	Factor 2
WISC	.84	
Pick the Picture	.85	
Simon Task		.99

Cumulative % of variance explained: 81%

Table 3

Test Statistics for Significant Effects of Age and Delay on Recognition Accuracy

Effect	*F*	*df*	*p*	η^2_p
Age	13.72	1, 186	<.001	.07
Delay	38.79	1, 186	<.001	.17
Age x Delay interaction	5.58	1, 186	.02	.03

Table 4

Standardized Coefficients and Test Statistics for Regressions with Age as a Predictor

Dependent Variable	Beta	*t*	*p*
Pick the Picture	.42	6.16	<.001
WISC	.60	9.95	<.001
Encoding	.61	10.61	<.001
Recognition	.25	3.51	<.001
Source Monitoring	.14	1.98	.05

Table 5

Standardized Coefficients and Test Statistics for Executive Function Predicting Recognition and Source Accuracy

Dependent Variable	Predictor	Beta	*t*	*p*	*p* when age is entered
Recognition	WISC	.42	5.01	<.001	<.001
	Pick the Picture	-.01	-0.13	.90	.93
	Simon	-.13	-1.69	.09	.09
Source Monitoring	WISC	.33	3.72	<.001	.001
	Pick the Picture	-.16	-1.84	.07	.09
	Simon	-.16	-1.98	.05	.05

Table 6

Standardized Coefficients and Test Statistics for Working Memory as a Mediator between Age and Source Monitoring

Variables Entered	Beta	*t*	*p*
Step 1: Age on source monitoring	.14	1.98	.05
Step 2: Age on WISC	.60	9.95	<.001
Step 3: WISC on source monitoring	.24	2.56	.01
Step 4: Age on source monitoring, controlling for WISC	-.03	-.37	.71

Table 7

Standardized Coefficients and Test Statistics for Encoding as a Mediator between Age and Source Monitoring

Variables Entered	Beta	*t*	*p*
Step 1: Age on source monitoring	.14	1.98	.05
Step 2: Age on encoding	.61	10.61	<.001
Step 3: Encoding on source monitoring	.33	3.73	<.001
Step 4: Age on source monitoring, controlling for encoding	-.06	-.66	.51

Table 8

Standardized Coefficients and Test Statistics for Encoding as a Mediator between Working Memory and Source Monitoring

Variables Entered	Beta	*t*	*p*
Step 1: WISC on source monitoring	.22	2.93	.004
Step 2: WISC on encoding	.54	8.54	<.001
Step 3: Encoding on source monitoring	.24	2.72	.007
Step 4: WISC on source monitoring, controlling for encoding	.09	1.03	.30

Table 9

Test Statistics for Significant Effects of Age, Delay and Training on Source Accuracy

Effect	*F*	*df*	*p*	η^2_p
Age	3.21	1, 182	.08	.02
Delay	7.07	1, 182	.01	.04
Training	5.04	1, 182	.03	.03
Three-way interaction	4.22	1, 182	.04	.02

Table 10

Sample Sizes within Each Cell for Study 2

Age Group	Control Condition	Training Condition
4-5 years	23	23
7-8 years	24	26

Table 11

Statistics for Outliers on BRIEF Subscales

BRIEF Subscale	Number of Outliers	Cutoff Score (Q3+1.5*IQR)	Outlier Scores
Inhibit	2	28	29
Shift	1	20	23
Emotional Control	4	26.5	27, 28
Initiate	1	20	22
Initiate –excluding items 10 and 66	2	16.5	17

Table 12

Reliability Estimates for BRIEF Subscales and Composite Scores

BRIEF Subscale	Number of Items	Cronbach's Alpha
Inhibit	10	.88
Shift	8	.74
Emotional Control	10	.81
Initiate	6 (2 items removed)	.71
Working Memory	10	.85
Plan/Organize	12	.86
Organization of Materials	6	.84
Monitor	8	.78
BRI	28	.88
MI	42	.94

Table 13

Factor Loadings of Subscale Scores on Two Factors

BRIEF Subscale	Factor 1	Factor 2
Inhibit		.83
Shift	.50	.53
Emotional Control		.86
Initiate	.73	
Working Memory	.76	
Plan/Organize	.90	
Organization of Materials	.73	
Monitor	.67	.51

Cumulative % of variance explained: 68%

Table 14

Descriptive Statistics for BRIEF Subscales and Composite Scores

BRIEF Subscale	N	Mean	Median	Maximum	SD	Range
Inhibit	73	16.29	15	30	4.66	19
Emotional Control	71	17.00	17	30	4.23	18
Initiate	72	12.73	13	18	2.85	14
Working Memory	75	17.07	17	30	3.90	17
Plan/Organize	75	19.11	19	36	4.64	19
Organization of Materials	75	11.64	12	18	3.05	12
Monitor	75	13.93	14	24	3.10	13
BRI	70	31.81	31	60	6.04	27
MI	72	74.48	74	126	14.61	63

Table 15

Mean Scores on BRI Subscales by Age Group

Age Group	Inhibit	Emotional Control
4-5 years	17.45 (4.83)	17. 56 (3.44)
7-8 years	14.70 (3.18)	15.54 (3.26)

Note: Standard deviations are in parentheses.

Table 16

Standardized Coefficients and Test Statistics for Age, MI and BRI Predicting Source Accuracy

Predictor	Beta	*t*	*p*	*p* when age is entered
BRI	-.36	-2.62	.01	.47
MI	.29	2.10	.04	.53
Age	.50	4.29	<.001	

Table 17

Mean Correct Recognition Proportions by Age Group and Condition

Age Group	Control	Training
4-5 years	.76 (.21)	.85 (.09)
7-8 years	.92 (.09)	.93 (.09)

Note: Standard deviations are in parentheses.

Table 18

Mean Proportions of Error Types by Age Group and Condition

Age Group	Interview Condition	Internal Intrusions	External Intrusions	Don't Know Responses
4-5 years	Control	.43 (.25)	.22 (.17)	.35 (.32)
	Training	.55 (23)	.18 (.16)	.27 (.28)
7-8 years	Control	.73 (17)	.13 (.11)	.15 (.16)
	Training	.63 (.23)	.11 (10)	.26 (.22)

Note: Standard deviations are in parentheses.

Figure 1. Distribution of Simon task accuracy scores.

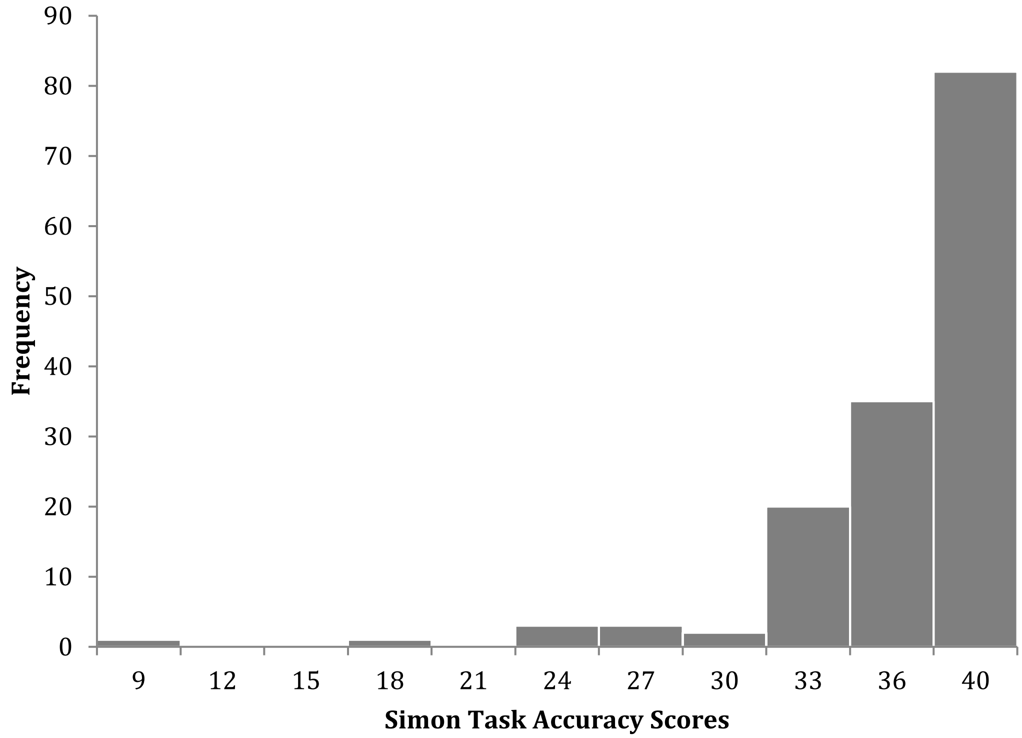


Figure 2. Distribution of Simon task reaction time differential.

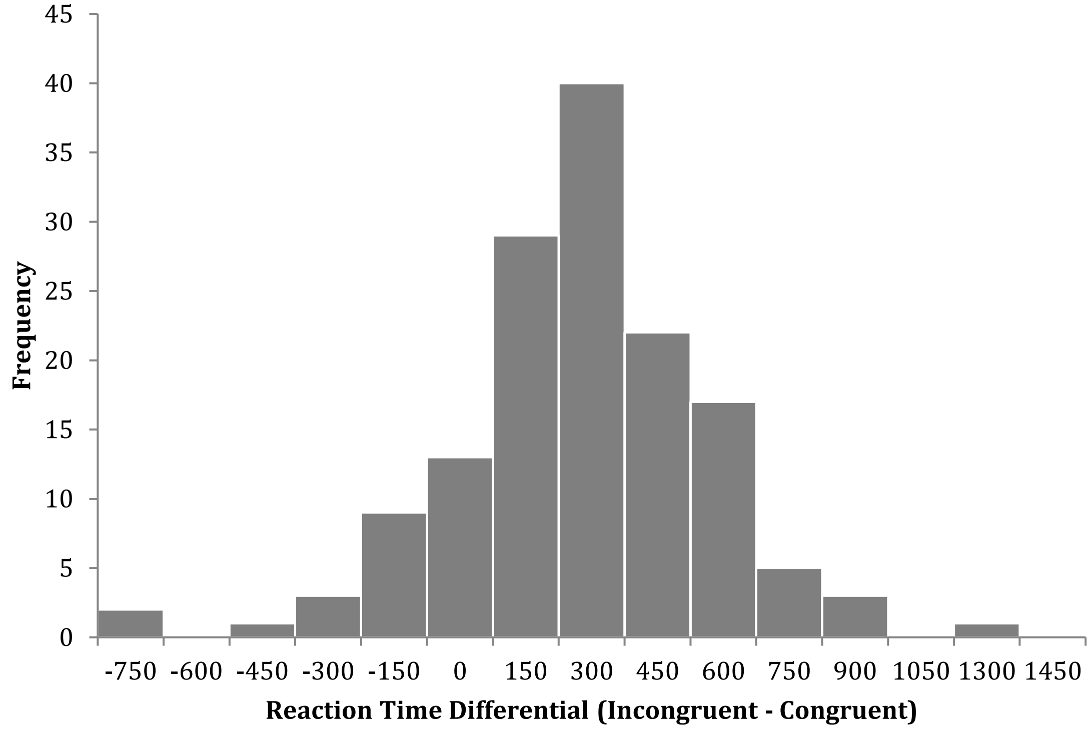


Figure 3. Distribution of Pick the Picture accuracy scores.

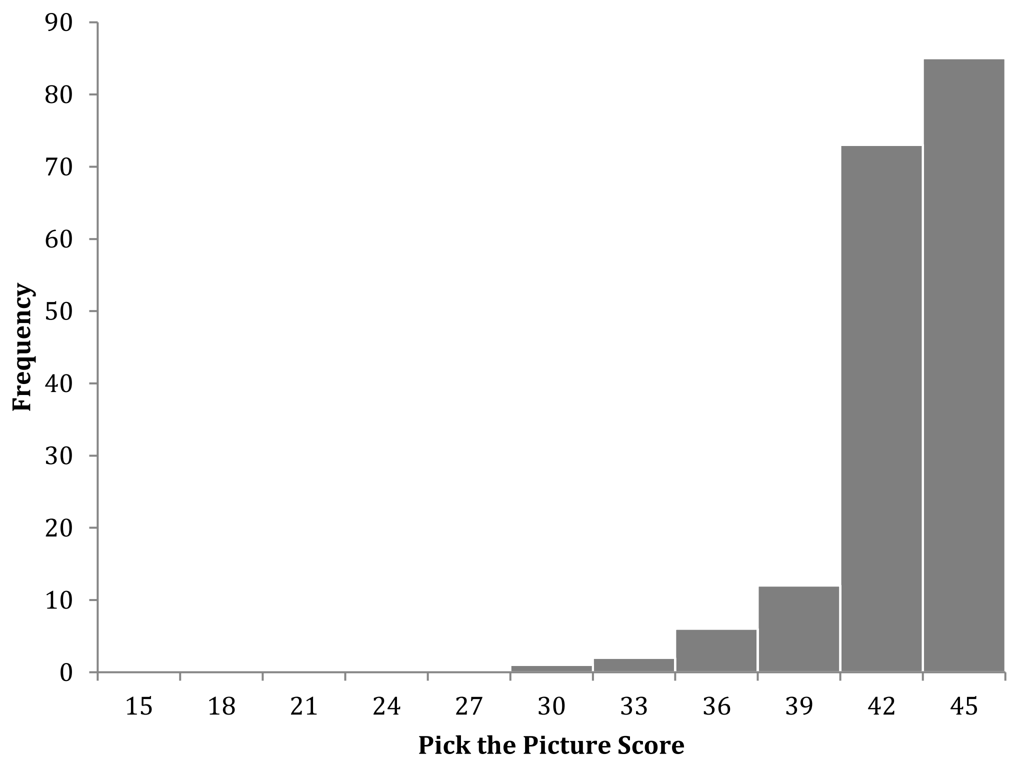


Figure 4. Distribution of WISC Digit Span scores.

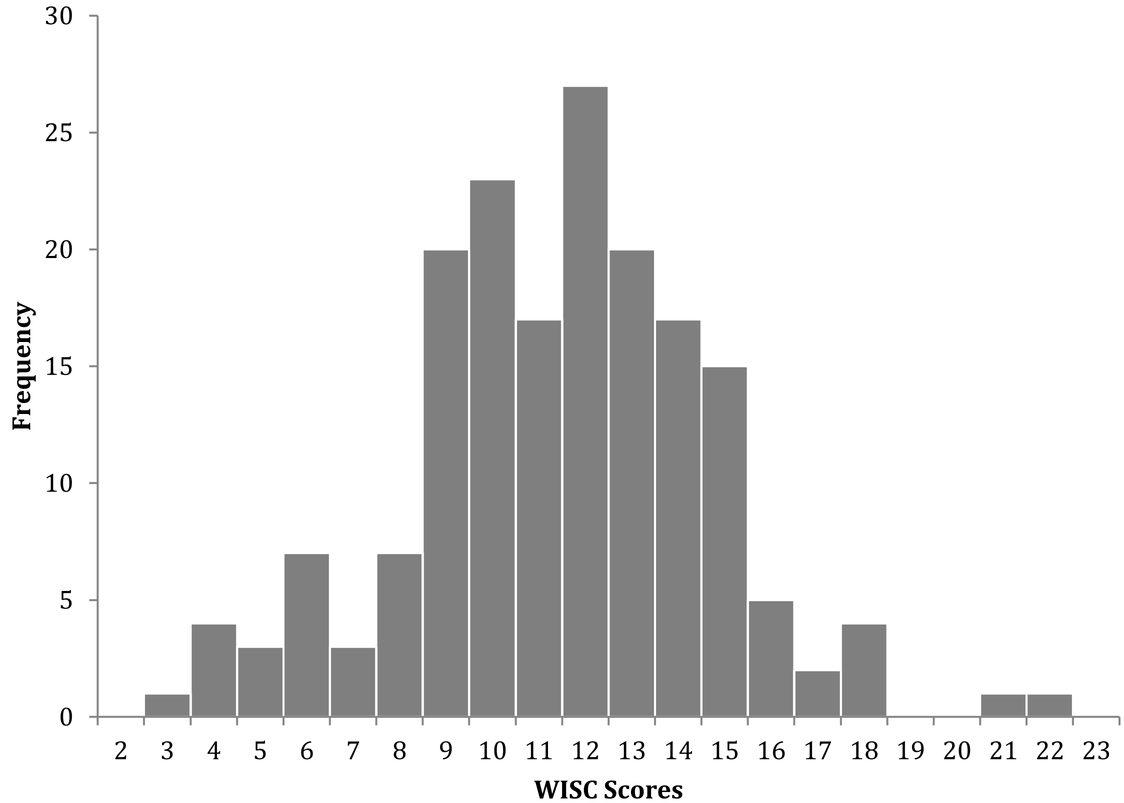


Figure 5. Working Memory as a Mediator between Age and Source Monitoring.

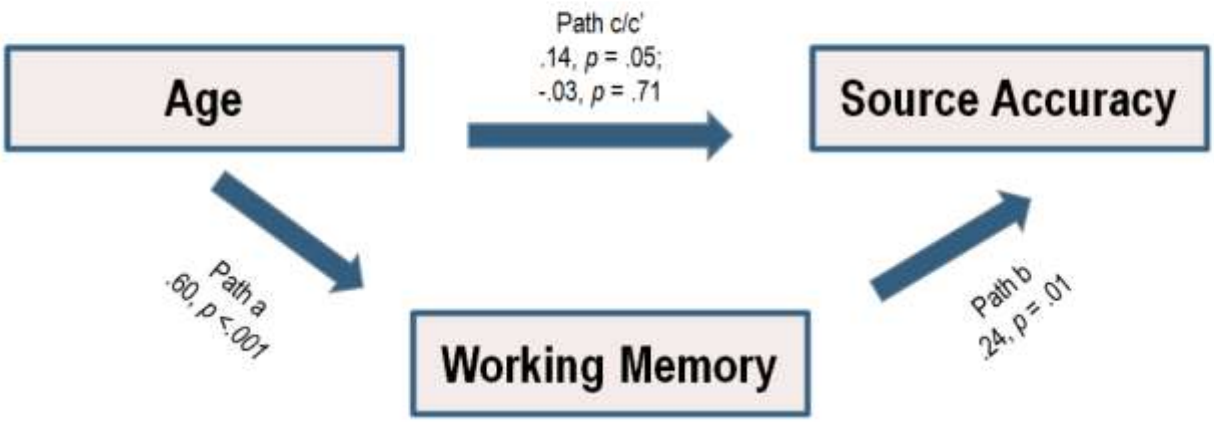


Figure 6. Encoding as a Mediator between Age and Source Monitoring.

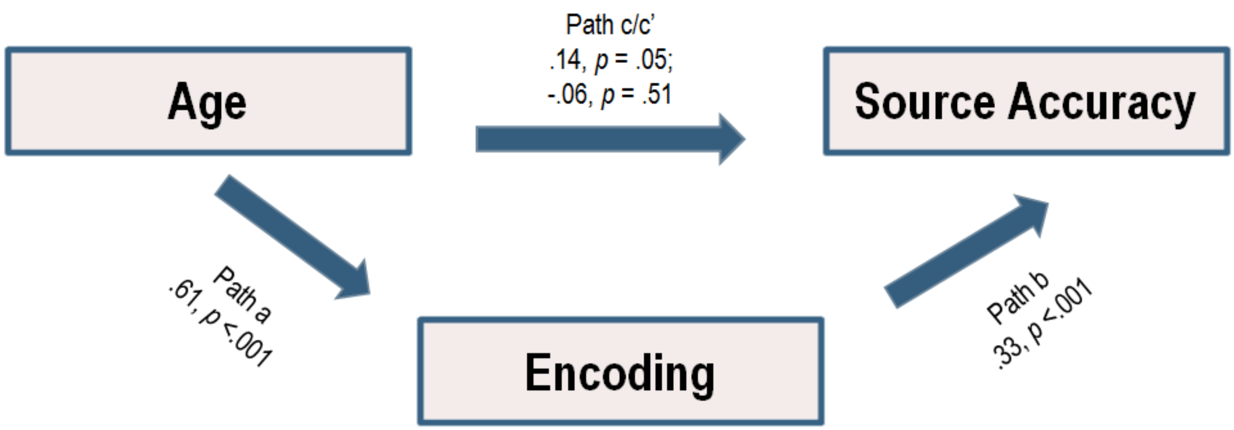


Figure 7. Encoding as a Mediator between Working Memory and Source Monitoring.

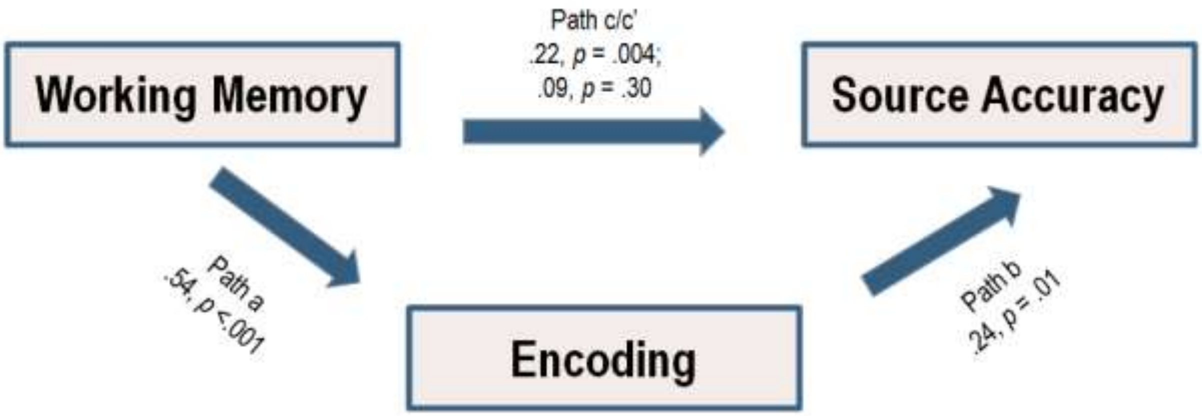


Figure 8. Distribution of BRI Scores.

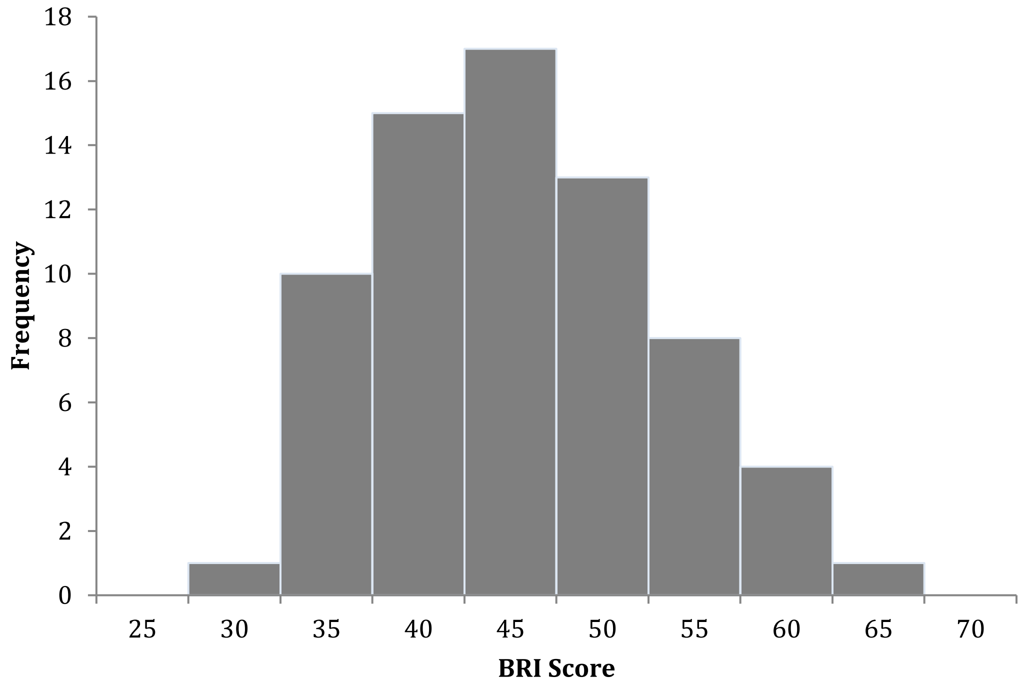
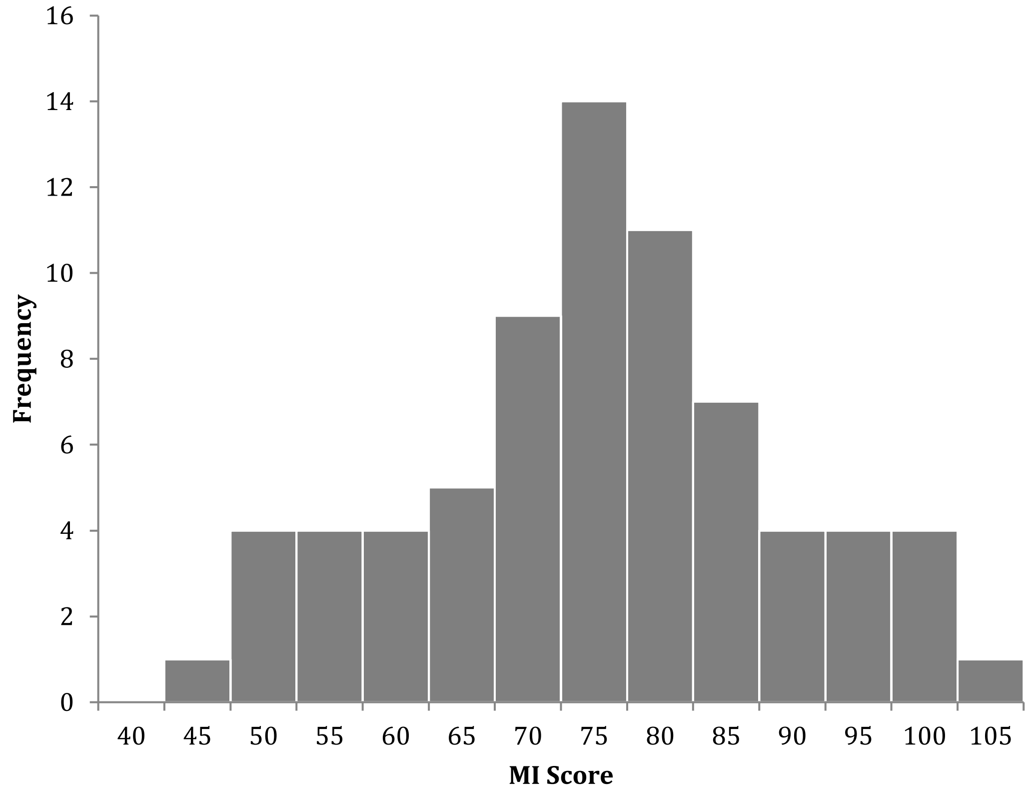


Figure 9. Distribution of MI Scores.



References

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

Baddeley, A. (1992). Working memory. *Science*, *255,* 556-559. doi: 10.1126/science.1736359

Baron, R.M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51,* 1173-1182.

Bjorklund, D. F., Miller, P. H., Coyle, T. R., & Slawinski, J. L. (1997). Instructing children to use memory strategies: Evidence of utilization deficiencies in memory training studies. *Developmental Review*, *17,* 411-441. doi: 10.1006/drev.1997.0440

Brainerd, C. J., & Reyna, V. F. (1995). Autosuggestibility in memory development. *Cognitive Psychology, 28,* 65–101.

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

Brainerd, C. J. & Reyna, V. F. (2004). Fuzzy-trace theory and memory development. *Developmental Review, 24,* 396-439.

Bright-Paul, A., Jarrold, C., & Wright, D.B. (2008). Theory-of-mind development influences suggestibility and source monitoring. *Developmental Psychology*, *44,* 1055-1068.

Brocki K. C, & Bohlin G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology, 26,* 571–593.

Brubacher, S. P., Glisic, U., Roberts, K. P., & Powell, M. (2011). Children's ability to recall unique aspects of one occurrence of a repeated event. *Applied Cognitive Psychology*, *25*, 351-358.

Bruck, M., & Ceci, S. J. (1995). Amicus brief for the case of State of New Jersey v. Michaels presented by Commitee of Concerned Social Scientists. *Psychology, Public Policy, & Law, 1*, 272-322. doi:10.1037/1076-8971.1.2.272

Bruck, M., Ceci, S.J., & Melnyk, L. (1997). External and internal sources of variation in the creation of false reports in children. *Learning and Individual Differences, 9,* 289-316.

Bruck, M., & Melnyk, L. (2004). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology, 18,* 947-996. doi:10.1002.acp.1070

Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology, 50,* 419-439. doi: 10.1146/annurev.psych.50.1.419

Connolly, D. A., & Lindsay, D. S. (2001). The influence of suggestions on children's reports of a unique experience versus an instance of a repeated experience. *Applied Cognitive Psychology, 15,* 205–223. doi:10.1002/1099-0720(200103

Craik, F.I. & McDowd, J.M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13,* 474-479.

Day, K., Howie, P. & Markham, R. (1998). The role of similarity in developmental differences in source monitoring. British Journal of Developmental Psychology, 16(2), 219-232. doi:10.1111/j.2044-835X.1998.tb00920.x

De Luca, C.R. & Leventer, R.J. (2008). Developmental trajectories of executive functions across the lifespan. In Anderson, V., Jacobs, R. & Anderson, P.J. *Executive functions and the fronal lobes: A lifespan perspective.* New York: Taylor & Francis, pp. 3-21.

Dunn, L.M., & Dunn, D.M. (2007). The Peabody Picture Vocabulary Test, Fourth Edition. Bloomington, MN: NCS Pearson, Inc.

Earhart, B., & Roberts, K. P. (2014). The role of executive function in children's source monitoring with varying retrieval strategies. *Frontiers in Psychology, 5*, 405. doi: 10.3389/fpsyg.2014.00405

Farrar, M. J., & Goodman, G. S. (1992). Developmental changes in event memory. *Child Development, 63*, 173-187.

Frost, P., Ingraham, M., & Wilson, B. (2002). Why misinformation is more likely to be recognised over time: A source monitoring account. *Memory, 10*, 179-185. doi:10.1080/09658210143000317.

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*(1), 51-60. doi:10.2307/1129860

Foley M. A., Harris J. F., Hermann S. (1994). Developmental comparisons of the ability to discriminate between memories for symbolic play enactments. *Developmental Psychology, 30,* 206–217. doi: 10.1037/0012-1649.30.2.206

Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin, 113*, 44-66.

Friedman, W. J. (2004). Time in autobiographical memory. *Social Cognition,22*, 591-605.

Frost, P.J. (2000).The quality of false memory over time: Is memory for misinformation "Remembered" or "Known"? *Psychonomic Bulletin and Review, 7,* 531–536.

Gerardi-Caulton, G. (2000). Sensitivity to spatial conflict and the development of self-regulation in children 24–36 months of age. *Developmental Science, 3*(4), 397-404. doi:10.1111/1467-7687.00134

Gerrie, M. P., & Garry, M. (2007). Individual differences in working memory capacity affect false memories for missing aspects of events. *Memory, 15*, 561-571. doi:10.1080/09658210701391634

Giles, J. W., Gopnik, A. & Heyman, G. D. (2002). Source monitoring reduces the suggestibility of preschool children. *Psychological Science, 13*, 288-291. doi: 10.1111/1467-9280.00453

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1-67.

Gioia, G., Isquith, P.K., Guy, S.C., and Kenworth, L.; Reviewed by Baron, I.S. (2000). Test review: Behavior rating inventory of executive function. *Child Neuropsychology 6,* 235–238. doi:10.1076/chin.6.3.235.3152

Guadagno, B. L., Powell, M. B. & Wright, R. (2006). Police officers and legal professionals' perceptions regarding how children are, and should be, questioned about repeated abuse. *Psychiatry, Psychology and Law, 13,* 251-260. doi:10.1375/pplt.13.2.251

Hutcheson, G. D., Baxter, J., Telfer, K., & Warden, D. (1995). Child witness statement quality: Question type and error of omission. *Law and Human Behavior Impact, 19,* 631. doi:10.1007/BF01499378

Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56,* 326-338.

Jaswal, V. K., Carrington Croft, A., Setia, A. R., & Cole, C. A. (2010). Young children have a specific, highly robust bias to trust testimony. *Psychological Science, 21*, 1541-1547.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S., (1993). Source monitoring. *Psychological Bulletin, 114,* 3-28. doi:10.1098/rstb.1997.0156

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88,* 67-85.

Karpinski, A. C., & Scullin, M. H. (2009). Suggestibility under pressure: Theory of mind, executive function, and suggestibility in preschoolers. *Journal of Applied Developmental Psychology, 30,* 749-763. doi:10.1016/j.appdev.2009.05.004

Kolb, D. A., Boyatzis, R. E., & Mainemelis, C. (2001). Experiential learning theory: Previous research and new directions. *Perspectives on thinking, learning, and cognitive styles.* (pp. 227-247) Lawrence Erlbaum Associates Publishers, Mahwah, NJ.

Koriat, A., & Goldsmith, M. (1996). Memory metaphors and the real-life/laboratory controversy: Correspondence versus storehouse conceptions of memory. *Behavioral and Brain Sciences, 19,* 167-188.

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

Lamb, M.E., Orbach, Y., Hershkowitz, I., Esplin, P. W., & Horrowitz, D. (2007). Structured forensic interview protocols improve the quality and informativeness of investigative interviews with children. A review of research using NICHD investigative interview protocol. *Child Abuse & Neglect, 31,* 1201-1231. doi:10.1016/j.chiabu.2007.03.021

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. 257-287). Mahwah, NJ: Erlbaum, Lindsay, D.S., Allen, B.P., Chan, J.C.K., & Dahl, L.C. (2004). Eyewitness suggestibility and source similarity: Intrusions of details from one event into memory reports of another event. *Journal of Memory and Language, 50,* 96–111.

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

Lloyd, M. E., Doydum, A. O., & Newcombe, N. S. (2009). Memory binding in early childhood: Evidence for a retrieval deficit. *Child Development, 80,* 1321-1328.

Mammarella, N. & Fairfield, B. (2008). Source monitoring: The importance of feature binding at encoding. *European Journal of Cognitive Psychology, 20,* 267-273.

Marsh, R. L., Landau, J. D., & Hicks, J.L. (1997). Contributions of inadequate source monitoring to unconscious plagiarism during idea generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 4,* 886-897. doi:10.1037/0278-7393.23.4.886

Martin-Rhee, M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition, 11*, 81-93.

Melinder, A., Endestad, T., & Magnussen, S. (2006). Relations between episodic memory, suggestibility, theory of mind, and cognitive inhibition in the preschool child. *Scandinavian Journal of Psychology, 47,* 485-495. doi:10.1111/j.1467-9450.2006.00542.x

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development, 13*, 395-423.

Nelson, K. (Ed.). (1986). *Event knowledge: Structure and function in development.* Hillsdale, NJ: Erlbaum.

Oates K., & Shrimpton, S. (1991). Children's memories for stressful and non-stressful events. *Medicine, Science and the Law, 31*, 4-10.

O'Neill, D. K., Astington, J. W., & Flavell, J. H. (1992). Young children's understanding of the role that sensory experiences play in knowledge acquisition. *Child Development, 63,* 474-490.

Ornstein, P. A., Gordon, B. N., & Larus, D. M. (1992). Children's memory for a personally experienced event: Implications for testimony. *Applied Cognitive Psychology, 6,* 49-60. doi:10.1002/acp.2350060103

Pearse, S. L., Powell, M. B., & Thomson, D. M. (2003). The effect of contextual cues on children's ability to remember an occurrence of a repeated event. *Legal and Criminological Psychology, 8,* 39–50.

Peterson, C., & Rideout, R. (1998). Memory for medical emergencies experienced by 1- and 2-year-olds. *Developmental Psychology, 34,* 1059-1072. doi: 10.1037/0012-1649.34.5.1059

Pipe, M. E., Lamb, M. E., Orbach Y. & Esplin, P. W. (2004). Recent research on children's testimony about experienced and witnessed events. *Developmental Review, 24,* 440-468.

Poole, D. A., & Lindsay, D. S. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology: Applied, 7,* 27-50. doi:10.1037/1076-898X.7.1.27

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

Powell, M. B., & Thomson, D. M. (2003). Improving children's recall of an occurrence of a repeated event: Is it a matter of helping them to generate options? *Law and Human Behavior, 27,* 365−384.

Powell, M. B., & Thomson, D. M. (1996). Children's memory of an occurrence of a repeated event: Effects of age, repetition, and retention interval across three question types. *Child Development, 67*, 1988-2004.

Raye, C.L. & Johson, M.K. (1980). Reality monitoring vs. discriminating between external sources of memories. *Bulletin of Psychonomic Society, 15*, 405-408. doi: 10.3758/BF03334572

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

Roberts, K.P., & Blades, M. (1996). Children's eyewitness testimony for real-life and fantasy events. In G.M. Stephenson & N.K. Clark (Eds.), *Investigative and Forensic Decision-Making. Issues in Criminological Psychology Series* (pp. 52-57). Leicester, UK: British Psychological Society.

Roberts, K.P., & Blades, M. (1999). Children's memory and source monitoring for real-life and televised events. *Journal of Applied Developmental Psychology, 20,* 575-596. doi:10.1016/S0193-3973(99)00030-1

Roberts, K.P., & Powell, M.B. (2005)a. Evidence of metacognitive awareness in children who have experienced a repeated event. *Applied Cognitive Psychology, 19,* 1019-1032. doi:10.1002/acp.1145

Roberts, K. P., & Powell, M. B. (2005)b. The relation between inhibitory control and children's eyewitness memory. *Applied Cognitive Psychology*, *19,* 1003-1018. doi: 10.1002/acp.1141

Roberts, K.P., & Powell, M. B. (2006). The consistency of false suggestions moderates children's reports of a single instance of a repeated event: Predicting increases and decreases in suggestibility. *Journal of Experimental Child Psychology, 94,* 68-89. doi:10 .1016/j/jecp.2005.12.003

Roberts, K. P., & Powell, M. B. (2001). Describing individual incidents of sexual abuse: A review of research on the effects of multiple sources of information on children's reports. *Child Abuse & Neglect, 25*, 1643-1659.

Robinson, E. J. (2000). Belief and disbelief: Children's assessments of the reliability of sources of knowledge about the world. In K. P. Roberts & M. Blades (Eds.), *Children's source monitoring* (pp. 59–84). Mahwah, NJ: Erlbaum.

Robinson, E. J., Butterfill, S. A., & Nurmsoo, E. (2011). Gaining knowledge via other minds: Children's flexible trust in others as sources of information. *British Journal of Developmental Psychology, 29*, 961-980.

Roebers, C. M., & Kauer, M. (2009). Motor and cognitive control in a normative sample of 7-year-olds. *Developmental Science, 12*, 175-181. doi:10.1111/j.1467-7687.2008.00755.x

Roebers, C. M., Moga, N., & Schneider, W. (2001). The role of accuracy motivation on children's and adults' event recall. *Journal of Experimental Child Psychology*, *78*, 313-329. doi:10.1006/jecp.2000.2577

Roebers, C. M., & Schneider, W. (2005). Individual differences in young children's susceptibility: Relations to event memory, language abilities, working memory, and executive functioning. *Cognitive Development, 20,* 427-447. doi:10.1016/j.cogdev.2005.05.006

Rosenthal, R. (1995). State of New Jersey v. Margaret Kelly Michaels case: An overview. *Psychology Public Policy and Law, 1,* 246-271. doi:10.1037/1076-8971.1.2.246

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

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

Scofield, J. & Behrend, D. A. (2008). Learning words from reliable and unreliable speakers. *Cognitive Development, 23,* 278-290. doi:10.1016/j.cogdev.2008.01.003

Tang, C. M. & Bartsch, K. (2012). Young children's recognition of how and when knowledge was acquired. *Journal of Cognition and Development, 13,* 372-394. doi: 10.1080/15248372.2011.577759

Thierry, K.L. (2009). Practice retrieving source enhances young children's discrimination of live and story events. *Journal of Applied Developmental Psychology, 30,* 552-562. doi: 10.1016/j.appdev.2008.12.030

Thierry, K. L., Lamb, M. E., Pipe, M.E., & Spence, M. J. (2010). The flexibility of source-monitoring training: Reducing young children's source confusions. *Applied Cognitive Psychology, 24,* 626-644. doi:10.1002/acp.1574

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

Thierry, K. L., Spence, M. J., & Memon, A. (2001). Before misinformation is encountered: Source monitoring decreases child witness suggestibility. *Journal of Cognition and Development, 2,* 1-26. doi: 10.1207/S15327647JCD0201_1

Titcomb, A. L., & Reyna, V. F. (1995). Memory interference and misinformation effects. In F. N. Dempster & C. J. Brainderd (Eds.), Interference and Inhibition in Cognition (pp.263-294). San Diego, CA: Academic Press.

Trocmé, N., Fallon, B., MacLaurin, B., Sinha, V., Black, T., Fast, E., Felstiner, C., Hélie, S., Turcotte, D., Weightman, P., Douglas, J., & Holroyd, J. (2010). "Chapter 1: Introduction." *Canadian incidence of reported child abuse and neglect – 2008: Major findings.* Ottawa.

Tulving, E. (1984). Precis of elements of episodic memory. *The Behavioral and Brain Sciences, 7,* 223-268.

Underwood, J. & Pezdek, K. (1998). Memory suggestibility as an example of the sleeper effect. *Psychonomic Bulletin & Review, 5,* 449-453.

Waters, G. M., & Beck, S. R. (2012). How should we question young children's understanding of aspectuality? *The British Journal of Developmental Psychology, 30,* 376-392. doi:http://dx.doi.org/10.1111/j.2044-835X.2011.02044.x

Wechsler, D. (2003). Wechsler Intelligence Scale for Children-Fourth Edition. Administration and scoring manual. San Antonio, TX: Harcourt Assessment, Inc.

Welch-Ross, M. K., Diecidue, K. & Miller, S.A. (1997). Young children's understanding of conflicting mental representation predicts suggestibility. *Developmental Psychology, 33,* 43-53. doi: 10.1037/0012-1649.33.1.43

Welch-Ross, M.K. (2000). A mental-state reasoning model of suggestibility and memory source monitoring. In K.P. Roberts & M. Blades (Eds). *Children's Source Monitoring* (pp. 227-255). Lawrence Erlbaum Associates Inc., New York, NY.

Wellman, H.M., Fang, F. & Peterson, C.C. (2011). Sequential progressions in a theory of mind scale: Longitudinal perspectives. *Child Development, 82,* 780-792.

Whitcombe, E. L., & Robinson, E. J. (2000). Children's decisions about what to believe and their ability to report the source of their belief. *Cognitive Development, 15,* 329-346.

Wiebe, S. A., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. A. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology, 108,* 436-452.

Williams, P. E., Weiss, L. G., & Rolfhus, E. L. (2003). WISC-IV technical report# 2: Psychometric properties. *The Psychological Corporation's WISC-IV Technical Manual, 2.*

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., & Tannock, R. (1999). Development of inhibitory control across the life span. *Developmental Psychology, 35,* 205-213.

Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment, 24,* 418-431. doi:10.1037/a0025779

Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2012). The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. *Psychological Assessment*, *24*, 226-239.

Willoughby M. T., Wirth R. J., Blair C. B., & Greenberg M. (2010). The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment*, *22,* 306–317.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood: VI. the development of executive function: Cognitive complexity and control--revised. *Monographs of the Society for Research in Child Development*, *68*, 93-119.

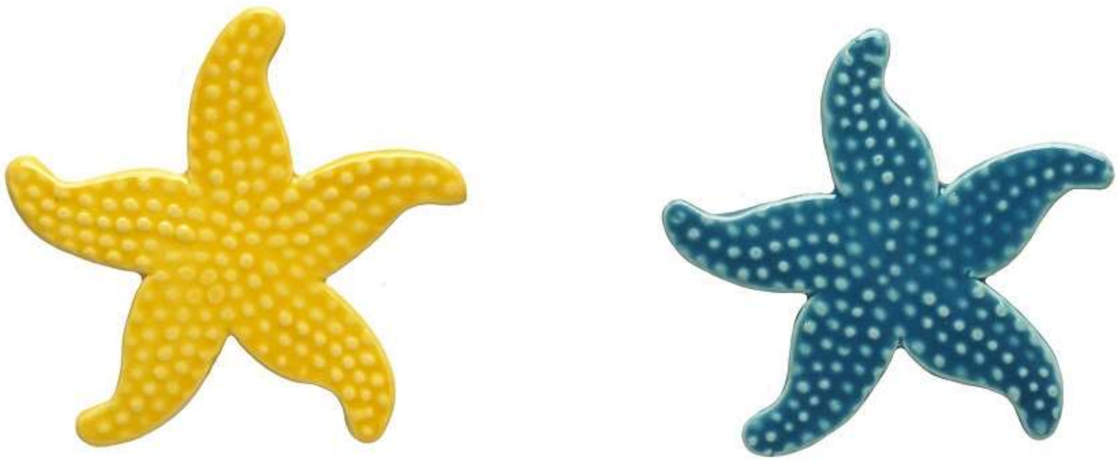
Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. *Blackwell handbook of childhood cognitive development.* (pp. 445-469) Blackwell Publishing, Malden. doi:10.1002/9780470996652.ch20

Appendix A: Interview Questions used in Study 1 Procedure

1. Did someone brush some teeth at HBA?
2. Was there a shovel at HBA?
3. Did someone hammer goldfish crackers at HBA?
4. Was there a picture of a whale at HBA?
5. Did someone use a banana as a phone at HBA?
6. Was there playdough at HBA?
7. Did someone touch a hot pot at HBA?
8. Was there dirt from the garden at HBA?
9. Did someone open a present at HBA?
10. Was there a sponge at HBA?
11. Did someone listen to their heart beat at HBA?
12. Was there cookies at HBA?
13. Did someone pat their head at HBA?
14. Was there a brain model at HBA?
15. Did someone play music at HBA?
16. Were there ribs at HBA?
17. Did someone kick a ball at HBA?
18. Did someone mix vinegar and baking soda at HBA?
19. Was Danny the Digester at HBA?
20. Was there a paper heart at HBA?
21. Did someone wear a helmet at HBA?
22. Were there apples at HBA?
23. Was there a bag of shiny stones at HBA?
24. Did someone pour dirty water through a filter at HBA?
25. Was there scissors at HBA?
26. Did someone tell a knock-knock joke at HBA?
27. Was there a red balloon at HBA?
28. Did someone put food colouring in water at HBA?
29. Was there paint at HBA?
30. Did someone feel toothpaste at HBA?
31. Was there orange juice at HBA?
32. Did someone jog on the spot at HBA?
33. Was there an organ apron at HBA?
34. Did someone play a computer game at HBA?
35. Was there a picture of a panda at HBA?
36. Did someone teach you about tonsils at HBA?

*Note: HBA stands for “Human Body Activities”. If children identified a detail, they were asked the follow up source question: “Did that happen in the real-life activity or the story?”

Appendix B: Simon Task Stimuli



Appendix C: Laurier Activities Event Script

1. Preparing the children for the Laurier Activities

- Gather the children.
- Say *"Hi my name is _______. Who knows the first letter of my name? "That's right. My name is ________ and the first letter of my name is __."*
- Tell them the following: *"I've brought you together to do something special with me today. We're going to do the Laurier Activities. Can you say that word for me again?"*..... (Children repeat *"Laurier Activities"*).
- Put up the **'L for Laurier'** somewhere just behind you so that the children can see it during the activities.
- Say *"The first thing we're going to do is sit down on **number squares."*** Hand out the **number squares** and instruct children to sit on them (one number square per child). Say *"When you get your number square, put it on the floor in front of me and sit on it."*
- Put on the **Red** cloak. Tell children *"There's only one Laurier cloak and I get to wear it because I'm the leader of the Laurier Activities. I'll tell you what we're going to do today"*.
- Tell the children that *"I was hoping that **Joe the Fox** would join us today in the Laurier activities but he just wants to say hello quickly because Joe's feeling very tired today. He says he's tired because **Mrs. Polar bear** kept him up all night. Have you seen a polar bear before?"*
- Put the **polar bear** next to the '**L for Laurier'** poster. *"I'll put the **polar bear** here next to the **'L for Laurier'** poster. Well the **polar bear** stayed over last night and she kept Joe up all night. Guess what she was doing?" (take children's guesses)* "***Scratching.*** *Can you make the sound of a polar bear scratching? Very annoying when you're trying to sleep. So you better say goodbye to Joe cause he's going to go back to sleep." (Bye Joe!)*

2. Introduce story

- Say *"Today's story is about **a dog in the city!**"*
- Say *"I wrote this story**!**"*
- Say *'I really like using bookmarks, so I'm going to use this bookmark with big **hearts.**'*
- "*So today's story is called **Ginger in the city**".*
- "*Now the story is all finished I'm going to put away my **hearts** bookmark."*

3. Puzzle time

- Say *"Now it's time to do a puzzle. There's only one puzzle and you all get to help me put it together. We've got to try to put the puzzle together so that it makes a funny clown. See if you can tell me what the clown is doing."* Children help RA put the puzzle together.
- *"Can everyone see the clown is **painting**?"*

4. Relaxation activity

- *"Now it's time to do the resting part of the Laurier Activities".*
- Say *"I'd like you all to lie down on your backs (legs stretched out straight) and close your eyes and keep them closed and just listen to me."*
- Open file labelled ***Ocean.***

- Read the following **very slowly** and calmly making sure that the children have their eyes closed and are quiet:

"I'd like you to keep your eyes closed and remain very calm and quiet now while we all rest... While we rest I'd like you all to think about the beach... Think about the ocean's waves rolling in… and out…
*As you are resting, think about what it would be like to be lying on the beach with the warm sun shining... I want you to try to relax your **legs**... think about how relaxed your legs feel when you hear how gently the waves roll… As you breathe calmly and slowly, think about the warm sun on your legs while you listen to the waves… Think about how warm and restful your **legs** feel as I come around and gently touch your **legs** to see if they're soft and warm."*

- Leader walks around to one child at a time touching their **legs** saying *"Do your **legs** feel soft and warm ______ (child's name)?"* Encourage children to respond with 'Yes'.
- Finish by saying *"Now keep your eyes closed while I count slowly to three. When I get to three, open your eyes and sit up. One....Two.....Three....."*

5. Getting refreshed
- Say *"The next thing to do during the Laurier activities is to make sure that you're all refreshed. It's important to feel refreshed after you've had a rest. Today you all get to refresh yourselves with **a fan**."* Leader fans each child with the electric fan.

6. Counting Objects
- Bring out ***flowers.***
- Say *"I brought some toy **flowers** with me today, but I am not sure how many I brought. Can you please help me count how many **flowers** I brought with me? (count the **flowers**)*
- Say *"Okay, great job. Now I'm going to put the **flowers** under this **blanket** that I brought. Then I am going to take some away and I want each of you to guess how many **flowers** are left under my **blanket**. So everyone close your eyes while I take some **flowers** away" (Let the children guess).*
- Once everyone has had a chance to guess, count the ***flowers*** again and say: *"Okay, well the Laurier Activities are almost over, so I am going to need your help putting the **flowers** away in this **hat** that I brought. Because I keep my **flowers** in a **hat** isn't that silly?"*

7. Packing up time and going back to classroom area
- Say *"Now it's time to pack up."*
- Say *"We have to pack up very quickly because I have to go **swimming**! I haven't gone yet this week, so it should be fun!"*
- *"Can you please give me your **number squares**?"*
- Say *"Who can remember what my name is?"* Let child answer. "That's right, you remembered that my name is _______."
- Say *"Well, we are all finished for today. I had a lot of fun. I hope you had fun too. Thank you very much for doing the Laurier activities with me today."*

Appendix D: Event Instantiations

No.	Item	A	B	C	D
1	Children sit on	Garbage bag	Carpet	Wash cloth	Number squares
2	Colour of cloak	Blue	Yellow	Green	Red
3	Noisy animal	Owl	Penguin	Seal	Polar Bear
4	Noise of animal	Laughing	Coughing	Yawning	Scratching
5	Source of story	Internet	Book store	Library	Leader wrote
6	Content of story	Party	Winter	Boat	Dog in City
7	Bookmark	Circles	Triangles	Squares	Hearts
8	Puzzle	Clown with umbrella	Clown juggling	Clown with balloons	Clown painting
9	Sound for relaxing	Birds	Thunderstorm	Heartbeat	Ocean
10	Part of body relaxed	Stomach	Nose	Arms	Legs
11	Getting refreshed	Baby wipes	Hand sanitizer	Drink water	Fan
12	Type of Object	Cars	Frogs	Shakers	Flowers
13	Put Objects Under	A Pillow Case	Umbrella	T-shirt	Blanket
14	Put Objects Away In	In a Lunchbox	In a cookie tin	In an Egg Carton	In a Hat
15	Next stop	Visit my grandma	Walking my dog	To a movie	Going swimming

Appendix E: Sample Items from the BRIEF

Inhibit: Has trouble putting the brakes on his/her actions.

Shift: Tries the same approach to a problem over and over even when it does not work.

Emotional Control: Overreacts to small problems.

Initiate: Has trouble coming up with ideas for what to do in play or free time.

Working Memory: When given three things to do, remembers only the first or the last.

Plan/Organize: Forgets to hand in homework, even when completed.

Organization of Materials: Cannot find things in room or school desk.

Monitor: Does not notice when his/her behaviour causes negative reactions.