2015

“Where Did I Learn That?” Exploring The Similarity Effect and Children’s Use of Memory Cues for Source Monitoring

Leanne E. Bird
Wilfrid Laurier University, bird4000@mylaurier.ca

“WHERE DID I LEARN THAT?” EXPLORING THE SIMILARITY EFFECT AND CHILDREN’S USE OF MEMORY CUES FOR SOURCE MONITORING

by

Leanne E. Bird

Bachelor of Arts, Honours Psychology

University of Guelph, 2013

THESIS

Submitted to the Department/Faculty of Psychology

in partial fulfilment of the requirements for

Master of Arts, Developmental Psychology

Wilfrid Laurier University

Abstract

An individual's ability to accurately monitor source (attribute known or remembered information to its particular *source* or origin) develops gradually throughout childhood. Along with task difficulty (i.e., delay between encoding and retrieval), *source similarity* is among the utmost hindrance to individuals' ability to accurately monitor source; specifically, the greater the similarity between sources the more difficult source monitoring judgments have been found to be, and the smaller similarity between sources (i.e., the greater number of differences between sources) the more accurate source monitoring judgments have been found to be. The *similarity effect* has been said to apply to all age groups, and has been assumed to be especially detrimental for young children. The present research looks further into the issue of source similarity, and suggests that the similarity effect may not be as generalizable as claimed. Specifically, although adults benefit most from dissimilar sources (as the similarity effect predicts), what may be paramount for young children (rather than *more differences between sources)* are few (at least one) but *distinct* differences between sources. The present study aims to begin consideration in this area by focusing on visual information. An experimental research design was used to assign 99 participants of different age groups (3-5, 6-8, 18-21) each to two different source-monitoring conditions. Each condition contained two actors, and the number of visual cues that differed between actors varied for each of the conditions (*one-cue* and *five-cues*). Specifically, the number of visual cues was manipulated such that one pair of actors displayed one distinct visual difference, and the other pair displayed five visual differences. After a short distractor task, participants were interviewed and asked to make source-monitoring judgments about actions performed by the actors within each of the two events. Data were collapsed and analyzed by age group. In line with past literature, an overall/general developmental progression was found to exist in participants' ability to make accurate source judgments. Contrary to the present proposed theory, there was no significant interaction between age and cue condition; individuals of all age groups were found to be more accurate in the five-cue condition than in the one-cue condition (as predicted by the similarity effect). The obtained results in relation to the proposed theory were discussed. Understanding the way that individuals use cues to monitor source can help us further understand developmental differences in source monitoring, clarify the basic mechanisms involved, and highlight other aspects of children's memory development. In addition, basic research questions concerning the nature of children's source-monitoring errors may be particularly important to understanding the caveats surrounding forensic interviews with young children.

Acknowledgments

First and foremost, I would like to express my sincerest gratitude to my supervisor, Dr. Kim Roberts, for her guidance, encouragement, knowledge, and support throughout my Masters degree and the thesis process. I feel privileged to have had the opportunity to work with such a renowned Academic and individual, and without her dedicated involvement to myself and to my research this work would not have been possible. For this, I thank her from the bottom of my heart.

I would also like to express gratitude to my advisory committee, Dr. Tobias Krettenauer and Dr. Alexandra Gottardo. Dr. Krettenauer was my Developmental Psychology professor in the first year of my Masters. His teaching style and enthusiasm for the topic made a strong impression on me, and I will always carry positive memories of his classes and his leadership. Dr. Gottardo has continually been a friendly and welcoming mentor within the department, and is thought of extremely highly both by myself and by my student colleagues. The valuable points that she has risen in discussions of my work, her unmatched attention to detail, and her kind-hearted support during this process has been very much appreciated.

Even though I have not had the opportunity to work with Dr. Tsang previously, I thank her very much for taking the time to be a member of my examination committee. She is an extremely reputable and respected individual in the field, and I feel honoured to have her be a part of this process.

Getting through this thesis has required more than academic support, and I have many people to thank for their friendship, help, and encouragement. First, I would like to thank the members of the Child Memory Lab as it has been a pleasure working with each and every one of them. Particularly, I would like to thank Candice Sommers, Alexia Lambis, and Hannah Giles for their instrumental help with data collection and coding. I would also like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC), as well as the children, parents, teachers and principals at the schools that participated. Without their help, data collection would not have been possible. Finally, I would like to dedicate a special, heart-felt thank you to Becky Earhart. Her encouragement, knowledge, and beautiful presence has been a guiding force throughout this process, and the companionship that I have found in her has meant more to me during this time than an acknowledgements section could possibly describe. Her selfless dedication to others and the academic community is unparalleled, and she is bound to make an amazing professor one day. I feel thankful and blessed to call her a friend.

Last but not least, I would like to thank my family and loved ones for their support during this time, as well as the encouragement, love, and laughter they have contributed to my life. Whether in person or in spirit, they have been my true teachers, guides, and favourite people in the world, and without them none of this would have been possible. Thank you.

Table of Contents

List of Tables

List of Figures

"Where Did I Learn That?"

Exploring the Similarity Effect and Children's Use of Memory Cues For Source Monitoring

Source Monitoring refers to the ability to attribute known or remembered information to its particular *source* or origin, including various different attributes such as when or where an event occurred or how it was perceived (Johnson, Hashtroudi, & Lindsay, 1993). For example, the ability to recall whether information originated from speaker A or speaker B, the media modalities through which information was provided, and whether a piece of information was directly experienced or suggested are all different forms of source monitoring. The term "source" refers to the conditions under which a memory was acquired, and source monitoring is the process of making decisions about the source of a memory (Johnson et al., 1993). The ability to monitor source can have various implications for everyday life, as memory for such information can have an impact on our thoughts and our actions. For example, if we remember hearing that an important meeting was cancelled, it is vital that we remember the source of that information in order to both evaluate the credibility of the source and determine the validity of the information (i.e., was it unreliable office gossip or did it come from the chair of the meeting?). Similarly, memory for source can also contribute to our ability to exert control over our own opinions and beliefs. For example, if you remember that the source of a 'statistic' was a newsstand tabloid, you have valuable information for evaluating the credibility of the supposed fact (Johnson et al., 1993). Along with its significance and practical implications for everyday life, source monitoring can be of particular importance to the domain of forensic interviewing. As noted by Roberts

(2002), interviewing victims of and witnesses to crimes is particularly important in investigations where there is little physical evidence. In situations such as these, the main evidence (including identification of sources or origin of event details) may come from eyewitness accounts, and concern has been raised about the reliability of eyewitness testimony (Roberts, 2002). Research into the ability of individuals to accurately monitor source can shed light upon witnesses' (particularly child witnesses') capabilities, interviewers' judgments of witness credibility, and can help contribute to the development of appropriate interviewing protocol (Roberts, 2002). The aim of the present research was to better understand the ability of different individuals to monitor source, including potential differences in the ways that individuals of various age categories may go about doing so.

The Developmental Path of Source Monitoring

An individual's ability to accurately monitor source has been known to develop gradually throughout childhood, and marked improvements have been found to take place in the 3- to 8-year-old range (Roberts, 2002). The development of source-monitoring skills is far from linear, and children may gain competence at some types of source distinctions (e.g., memories of actions performed by the self vs. others) before others (e.g., memories of performed vs. imagined events); in other words, it is domain specific rather than domain general (Foley & Johnson, 1985; Roberts, 2002). As mentioned, children's source monitoring is also gradual rather than abrupt. Children as young as 3 years old may be able to appropriately rely on informative sources more than uninformative sources (e.g., knowing that someone who has looked at an object is better informed than someone who has not) yet may not be able to later explain how they came to know certain information (e.g., who the person was) until after the ages of 5 or 6, suggesting that explicit awareness of sources or the ability to reflect upon them does not occur until after the development of 'implicit' source monitoring skills (Roberts, 2002; Robinson, 2000). Children's ability to accurately monitor source may also be sensitive to the nature of the task, as 3- to 4-year-olds have been found to be able to distinguish memories of performed and pretended actions when tested nonverbally, but not when tested verbally (Roberts, 2002; Roberts & Blades, 1995). In the forensic arena, the level of source confusions may be dependent upon interview technique, as illustrated by children who have made fewer source errors when they were allowed to freely recall events versus when they were questioned specifically about individual event details (Roberts, 2002; Roberts & Blades 1998, 1999). With these factors in mind, it is clear that children's source-monitoring precision may be sensitive to their developmental standing, and judgments about their competence should be made only with such realities in mind (Roberts, 2002).

Source-Monitoring Theory

As may be gathered by a most basic understanding, the main process driving the ability to monitor source is memory—hence, the ability to attribute *known or remembered information* to its particular source or origin, or to remember where it came from. However, according to *Source-Monitoring Theory* (Johnson et al., 1993; Johnson & Raye, 1981) (a theory commonly used to explain many research findings on source monitoring), source is attributed through an examination of memory characteristics and through strategic decision-making (Roberts, 2002).

The two central tenets of source-monitoring theory, the *examination of memory* *characteristics* and the *use of strategic decision-making* are each valuable in their own right. The former, examination of memory characteristics, can provide valuable cues to source as particular characteristics of memories can prompt recognition of the source (Roberts, 2002). For example, examining the sound of a person's voice in a memory can prompt recognition of the particular person who carried out the action. In addition, in contrast to memories of non-perceived events, memories of events that were actually perceived contain more perceptual, sensory, contextual, affective, and semantic information. Examination of the characteristics of such memories can help indicate that the event did in fact occur as well as aid the individual in making a source decision (Roberts, 2002). It is important to note that, in the case of examining memory characteristics, decisions about source are often carried out automatically with minimal strategy such as by using general knowledge or common sense to infer source. This automatic judgment process may involve quick decisions that occur in the course of remembering without conscious awareness of the decision-making process, and are normally based on qualitative or other such characteristics that were encoded when the memory itself was formed (e.g., that the clown at the party told you the joke, and not your friend). They also may be made on the basis of a match between the characteristics of a memory and knowledge about particular sources (Ferguson, Hashtroudi, & Johnson, 1992).

The latter tenet, strategic decision-making, is the second strategy that can be used to make a source decision, simultaneously or in addition to examination of memory characteristics (Roberts, 2002). This is a more analytic and deliberate way of attributing the source of a memory, and involves a controlled, or systematic, judgment process.

When monitoring source in such a way, an individual will reason carefully about what is possible given the information that they have from the memory, and employ strategies (such as retrieving supporting memories, reasoning about constraints, and using logical analysis) in order to arrive at a source decision (Johnson et al., 1993). For example, when remembering who gave you directions, you might recall where you were when you were given the information and reason that it must have been a person who was also in attendance who gave you the directions. Similarly, one might correctly attribute the memory of a conversation with a particular person to imagination, as they know that they have never been acquainted with that person (Ferguson et al., 1992).

According to source monitoring theory, there are at least two ways that source monitoring can fail. Firstly, the various memory characteristics may not be available or may not be salient; in other words, 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 (Ferguson et al., 1992). For example, an individual may fail to encode particular characteristics, or may encode the information but not access it in attempts at retrieval (Ferguson et al., 1992). Secondly, an individual may fail to use successful reasoning, strategies, or logic, or may engage in faulty forms of such (Ferguson et al., 1992). Overall, although source-monitoring theory does not provide an explicit outline of cognitive or developmental mechanisms, due to the examination of memory characteristics and strategic decision-making processes, the theory holds that source is inferred at the time of recollection rather than encoded at the time of the event (Roberts, 2002).

Memory Characteristics and Cues to Source

Although the use of strategy when making a source decision is for the most part straightforward and logical, it is less clear how different characteristics of memory are examined and made use of by individuals, and the functions of such characteristics in the process of source monitoring. Such discussion has led researchers in the field to analyze the roles of different characteristics. According to Ferguson and colleagues (1992), among the most valuable memory characteristics are records of contextual (spatial, temporal) information, semantic detail, affective information (e.g., emotional reactions), cognitive operations (e.g., imagining or elaborating), and perceptual information (e.g., colour, sound) that took place when the memory was formed. Decisions regarding source often involve evaluating the kind and number of these characteristics, as particular types of memories have been found to contain more or less of the particular characteristics. For example, as opposed to memories for internally generated events, memories for externally derived events have been found to include less information about cognitive operations, and more perceptual, contextual, semantic, and affective information (Ferguson et al., 1992). *Visual cues* (another type of memory characteristic and a form of perceptual information) are also among the most advantageous of characteristics as, if successfully encoded, they may be able to provide valuable indications of source (Ferguson et al., 1992). For example, when visual cues are encoded from two separate sources (such as one speaker being male, wearing a blue shirt and a baseball cap, while another being female, wearing a pink shirt and no baseball cap), they may be accessed during retrieval and assist the individual in making an accurate source judgment. One of the major benefits of visual memory information is that (if encoded effectively) it possesses the ability to *visually* distinguish one source from another, and provides information that need not rely as much on context, semantic detail, affective information, cognitive operations, or other perceptual information at the time of encoding.

The Nature of Children's Source Monitoring Errors

Cognitive Factors. Research on adults with frontal lobe damage, who present with many of the same problems that young children demonstrate in source monitoring, has suggested that the frontal lobe may be implicated in the development of source monitoring (Schacter, Kagan, & Leichtman, 1995). Differences in *Executive function,* a broad category of skills that support goal-directed behavior and that underlies many cognitive abilities (Earhart & Roberts, 2014), have been linked with immature frontal lobe development, broadly suggesting that executive function may play a role in accounting for developmental differences in source-monitoring ability (De Luca & Leventer, 2008). The two specific components of executive function that have been suggested to relate to source-monitoring accuracy are *inhibitory control* and *working memory*. Specifically, inhibitory control may relate to source monitoring due to its ability to inhibit familiarity-based retrieval processes as well as information from competing sources. For example, in addition to showing problems with inhibitory control tasks, adult patients with frontal lobe lesions showed a deficit in source monitoring (Luria, Pribram & Homskaya, 1964; Drewe, 1975). Correspondingly, inhibitory control was found to account for a significant proportion of variance in source suggestibility among 5- to 7-year-olds (Roberts & Powell, 2005). Working memory may relate to source monitoring due to its involvement in controlling attention, recognition memory, and playing a role in designating what information cognitive resources will be allotted to

(Earhart & Roberts, 2014; Gerrie and Garry, 2007; Ruffman et al., 2001). Both of these skills develop throughout childhood, and simultaneous improvements exist with source monitoring ability. 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 (Earhart & Roberts, 2014).

Other developmental processes including *theory of mind* (i.e., the ability to attribute mental states — beliefs, intents, desires, pretending, knowledge, etc. — to oneself and others, and to understand that others have beliefs, desires, and intentions that are different from one's own) (Premack & Woodruff, 1978) and reasoning about conflicting mental representations have also been shown to account for variance in source monitoring, and improvements in such areas are generally linked to age; in fact, executive function itself has been tied to age (Earhart & Roberts, 2014). As such, it could be argued that a more general "cognitive development factor" may be a stronger predictor of source monitoring accuracy than executive function alone (Bright- Paul et al., 2008; Earhart & Roberts, 2014; Welch-Ross et al., 1997; Welch-Ross, 1999).

Finally, referred to as the process of binding or *cohesion*, the features that comprise a given memory episode are not stored in a random manner; rather, they must be bound together so that they form a unique representation of the event (Lorsbach & Reimer, 2005). As children's ability to monitor source develops with age, in turn, so does their ability to bind source with content (i.e., the details of a memory; Lorsbach & Reimer, 2005). For example, Lorsbach and Reimer (2005) examined whether developmental differences exist in the ability to bind features together in a working memory task. They found that sixth-grade students experienced greater difficulty than college students, and third-grade students performed even worse than sixth-grade students in both their memory for individual features as well as their memory for the combination of those features (Lorsbach & Reimer, 2005). The memory-binding process has been considered crucial for the explicit memory system and, without adequate implementation, memory may be compromised and source errors may be made (Metcalfe, Mencl, & Cottrell, 1994). For example, fragmentary information of an episode may be remembered without a cohesive memory of where and when the information was acquired (Schacter, Norman, & Koustaal, 1998).

Shortcomings for Source Monitoring. Each of the processes outlined above play a role in individuals' ability to monitor source, and may have direct implications for examination of memory characteristics and strategic decision-making; the two central tenants of source monitoring theory. In turn, as they all tend to develop with age, young children may not have the resources necessary to combine multiple cues in a meaningful way that will benefit them in making accurate source judgments in the future (Lorsbach & Reimer, 2005). The fact that young children may be deficient in the kind of cognitive flexibility that is needed to use multiple cues has been alluded to in past source-monitoring literature, and it has been generally accepted that this lack of cognitive flexibility may underlie many of their failures in source identification (Foley, Wilder, McCall, & Van Vorst, 1993; Roberts, 2002).

Although many important processes develop with age, is it reasonable to dismiss young children as not possessing the ability to monitor source altogether, simply because they may not yet have the cognitive capacity to make use of multiple cues? Evidence from developmental studies of source monitoring have found that young children have performed as well as adults in some source-monitoring situations but not in others, indicating that their ability to *make accurate source judgments in general* may not simply come with age (Johnson et al., 1993). As such, if source monitoring itself is not a general ability that develops with age, it is important to consider the potential ways in which young children may differ from older children and adults in the way they use cues to monitor source.

The Issue of Similarity

The issue of *source similarity* has long been recognized in the literature. Along with task difficulty (i.e., delay between encoding and retrieval), source similarity has been found to be among the utmost hindrance to individuals' ability to accurately monitor source. According to the *similarity effect*, the more similar the sources are (i.e., the fewer differences between sources) the more difficult source monitoring judgments have been found to be, and the less similar the sources are (i.e., the greater number of differences between sources) the more accurate source monitoring judgments have been found to be (Roberts, 2002). Fundamentally, individuals are *more likely* to confuse memories of sources that are more similar than of those that are more different (Roberts, 2002). The similarity effect has been reproduced time and time again within adult populations (Roberts, 2002), and has been known to apply to young children as well. For example, children have been found to be more confused when asked to make source judgments about similar actions than dissimilar actions (Day, Howie, & Markham, 1998; Roberts & Blades, 1999), and children had more difficulty distinguishing between words spoken by two speakers of the same gender than words spoken by two speakers of opposite genders (Lindsay, Johnson & Kwon, 1991, Experiment 1).

Limitations of Current Literature. Further inquiry into the realm of source similarity points to limitations in the current body of academic and empirical literature. Although various studies have been conducted in the area of source similarity and conclusions have been gathered emphasizing the disadvantages of such circumstances, the majority of research concerning children has focused on the impact of source similarity within the realm of *reality monitoring* (i.e., monitoring of source for real vs. imagined events). For example, research has focused on the ways in which high similarity may cause individuals to confuse video or television events with reality (Thierry & Pipe, 2009), memories of words they had actually said aloud and memories of words they had imagined saying (Foley, Johnson, & Raye, 1983), memories of actions they imagined themselves performing and memories of actions they actually performed (Foley & Johnson, 1985; Foley, Johnson & Raye, 1983), and more. As mentioned previously, the theoretical foundation of many of these studies was likely based on the notion that perceptually based memories contain more contextual information, sensory attributes, vividness and detail whereas memories of imagined events contain more indications of the cognitive operations active at the time of the experience, and source monitoring decisions made on the basis of these characteristics become increasingly difficult as they become increasingly similar. Although studies such as these may be highly reputable, informative, and crucial for the understanding of source monitoring in general, one potentially negative outcome has been observed; that is, the current burdens of the similarity effect have been generalized and applied to the entirety of source monitoring, claiming that source monitoring errors will occur most often when the information in memories based on different sources is similar—particularly for children, on any dimension (Day et al., 1998).

Basis of the Present Research. As the similarity effect has been replicated in the literature and found to exist in a wide range of age groups, for what reason may generalization of the construct be negative? There is a very good reason to ask this question, which forms the basis and foundation of the present study. Essentially, at this point, the majority of studies manipulating source similarity have excluded young preschoolers as participants and have been in the realm of *reality monitoring* as opposed to a wide variety of dimensions, situations, or contexts (Thierry & Pipe, 2009). As a result of this, although it may be the case in the studies mentioned or undertaken, it is not clear that the similarity effect is a phenomenon that applies to all source monitoring situations, and a general explanation or rationalization of the similarity effect as applicable to individuals of *all* age groups in *all* source monitoring situations may be inappropriate.

The basis for this claim is grounded in two main motivations. First, in one of the most renowned and well-known studies of source monitoring in children, researchers manipulated the auditory similarity of a tape-recorded list of words (Lindsay et al., 1991, Experiment 1). Half of the words were presented from a speaker on the subject's left and the other half from a speaker on the subject's right, and for half of the subjects in each age group the same person's voice came from both speakers whereas for the remaining subjects a different voice (one male and one female) came from each speaker (Lindsay et al., 1991, Experiment 1). Participants were later asked to remember the source (left or right speaker) of particular words (Lindsay et al., 1991, Experiment 1). Although young children made significantly more accurate source monitoring judgments when the words were presented by male and female voices (in this case, the *more different* case), gender was indeed the *only* difference that existed between sources; there were no other visual, contextual, perceptual, or other, details. As opposed to other situations in which there may be various additional relevant cues available in memory to aid in making a source decision, this particular instance offered only one differentiating cue (which is a very small number). What is important to note about this study is that all participants (both young children and adults) performed exceptionally well in a situation in which only one distinct difference existed between sources (Lindsay et al., 1991, Experiment 1). As such, and opposed to what would be predicted by the central tenet of the similarity effect, it is not necessarily the case that children perform poorly when only a small number of differences exist between sources. This can be further emphasized by studies of repeated events in which multiple differences exist, yet children have difficulty tagging specific details to correct occurrence in a series of events (Powell & Thomson, 1996; Powell et al., 1999). Perhaps what children benefit from, instead of the *number of cues*, are cues that they find *distinct*.

In terms of the second motivation, it is clear from our discussion on the nature of children's source monitoring errors that children do not have the same cognitive flexibility as adults do, as processes such as inhibitory control, working memory, theory of mind, feature-binding or cohesion, and other such cognitive developmental factors critical to source monitoring tend to develop with age. Young children may not have the resources necessary to combine multiple cues in a meaningful way that will benefit them in making accurate source judgments in the future. This could account for many of their failures in source identification in general. Accordingly, it is possible to speculate that, of the sources that they are able to identify, young children may actually fair better with a few cues or a single distinctive cue than with a greater number of differences.

Should the proposed theory stand true and distinctiveness trump the *number of differentiating cues* for young children's source monitoring performance, the most important consideration will be to determine the various or potential cues that young children find beneficial and distinct. As illustrated in the study by Lindsay and colleagues (1991, Experiment 1), gender presents as the first candidate. Alongside gender, it is suggested that visual cues may be of particular importance. Although adults and older children may have the capacity to hold the multiple cues in working memory while simultaneously analyzing characteristics of the memory and using strategic logic/decision-making processes to make a source decision, young children with limited cognitive capacity may not have the working memory capacity necessary to engage in such types of binding exercises. As such, young children may be able to make better use of a small number of isolated visual details without putting too much demand on their resources or requiring them to engage in strategy or logic for which they have not yet acquired the flexibility. Conversation around the distinctiveness of visual cues is what the present research attempts to stimulate.

The Present Study

The objective of the proposed research was to look further into the issue of source similarity, to determine whether children benefit more from a single distinctive visual cue than a greater number of cues, and to better understand the developmental differences in the effect of cues on children's source-monitoring performance. A within-subjects experimental research design was used to expose 99 participants of different age groups (3-5, 6-8, 18-21 years) to two different source-monitoring events. Each event/condition contained two sources[1]. In order to compare single item differences to multiple item differences, a *one-cue* condition and a *five-cues* condition were created. Specifically, each condition contained two sources, and the number of distinguishing source cues contained within each pair of sources was manipulated such that one pair contained one salient difference (the one-cue condition), and the other contained five differences (the five-cues condition). After a short distractor task, participants were asked to make source-monitoring judgments in regards to each of the events, and a mean source-monitoring performance score was calculated for each of the age groups in each of the experimental conditions. In accordance with existing literature on source-similarity and young children's cognitive processing abilities, it was hypothesized that the following differences would exist:

Age Differences.

Hypothesis 1: There would be an overall/general developmental progression in participants' ability to make accurate source judgments.

Condition Differences.

Based on the rationale for the proposed study, it was further hypothesized that:

---

[1] The study utilized *people* as the form of source, as past research has found this to be an effective way to isolate and control the construct and its components.

Hypothesis 2: There would be a significant interaction between age and cue condition. Specifically,

- According to the similarity effect, adults would be more accurate in the five-cue condition than in the one-cue condition because the greater number of differences between sources the less similar they are, and therefore the easier they are to distinguish.
- Regarding the 3-5-year-olds, source scores may not follow the same pattern predicted by the similarity effect. Due to young children's limited cognitive capacity, even though there are a greater number of differences in the five-cue condition than the one-cue condition, preschoolers may not have the prerequisite cognitive development skills (e.g., working memory, inhibitory control) necessary to use multiple cues effectively. As such, 3-5-year-olds may find a single visual cue more distinctive than a greater number of differences (the opposite pattern than that predicted for adults) and, thus, perform more accurately in the one-cue condition than in the five-cue condition.

Data collected from the 6-8-year-olds was exploratory, as it was unclear whether and how their scores would differ in the one and five cue conditions. Therefore, no specific prediction was made regarding this age group.

Method

Participants

Ninety-nine participants from the local area were recruited to participate in the study. Thirty participants were between the ages of 3 and 5 years old, 37 participants were between the ages of 6 and 8 years old, and 32 participants were between the ages of 18 and 21 years old. See Table 1. Fifty-nine percent were female and 41 percent were male (see Table 2), and reported being from a wide variety of ethnic backgrounds (see Table 3). It was required that all participants be able to communicate in English, have normal development, and (if under the age of consent) obtain consent from a primary caregiver. Three participants were excluded from analyses, as their responses to study measures indicated that they did not understand the task at hand or demonstrated a response bias.

The time requirement was approximately 25-30 minutes on one occasion, which comprised time to complete all tasks.

Participants were recruited using the following strategies: 1) distribution of consent forms within the Waterloo Region District School Board, 2) utilization of the Wilfrid Laurier University undergraduate PREP system[2], 3) recruitment posters (hung in establishments such as grocery stores, recreation centers, and libraries), 4) advertisements (on internet sites such as the Child Memory Lab website), 5) snowball sampling, and 6) ‘word-of-mouth’ recruiting.

Individuals who signed up and participated by means of the Wilfrid Laurier University undergraduate PREP system were awarded 0.5 of a course credit for

---

[2] The Psychology Research Experience Program (PREP) offers students the opportunity to earn course credit by participating in psychology research studies.

participation in the study. Schools from the Waterloo Region District School Board that agreed to participate received a financial donation of $5 per signed parental consent form received (regardless of whether or not the student met selection criteria). Participants from the community willing to travel to the Child Memory Lab to participate were eligible to receive full parking reimbursement (if applicable) and $10 compensation, as well as a certificate of participation and a small prize (value ~$5) for child participants. All participants were advised that participation in the study was voluntary, and that they were free to decline to participate without penalty. In addition, should they decide to participate, they may withdraw from the study at any time. If they chose to end their participation before completing all study tasks they would still receive full compensation/remuneration. Participants were also advised that should they withdraw from the study before data collection was completed, their data (including videotapes) would be returned to them or destroyed, and if any participants withdrew from the study early, their data would not be transcribed or used in any publications about the study. Participants had the right to omit any question(s)/procedure(s) they chose without penalty or loss of benefits to which they were otherwise entitled. The study received full ethical approval from Research Ethics Boards governing both the Waterloo Region District School Board[3] and Wilfrid Laurier University[4].

Materials

Source-Monitoring Simulations. The study utilized a technique called the *cue salience technique*—a method of systematically varying the availability of particular

[3] Approved following the Research Committee meeting held on February 14, 2014

[4] REB File number 3937

memory characteristics (in this case, visual cues) in order to determine the effect of the manipulation and, therefore, their role in source-monitoring decisions (Ferguson et al., 1992). Participants watched a short video with two segments. Each segment displayed two sources (actors) performing a total of 12 actions, presented in a series of individual clips. Specifically, the first segment displayed clips of one actor performing the first six actions, followed by clips of the other actor performing the second six actions. After a short delay, the second segment began and exhibited the same cycle as the first, displaying clips of one actor performing the first six actions followed by clips of the other actor performing the second six actions, only this time with a different pair of actors than were presented in the first segment. The purpose of each action was to determine a target source for future source monitoring examination. All aspects of the clips within the video were controlled and identical in almost every way; for either segment, the only differences that existed between clips (aside from the actual actions performed) were the number of visual cues that differed between sources. Specifically, one segment contained one salient difference between the actors performing the first six and the second six actions (i.e., t-shirt colour), and the other segment contained five differences between actors performing the first six and the second six actions (i.e., t-shirt colour, hair colour, wearing/not wearing a hat, wearing/not wearing a necklace, colour of pants). See Appendix A. The segments of the video, in turn, constituted the two experimental conditions.

The clips in each condition were based on four action lists that were predetermined and created by the researcher. Each action list consisted of six actions that would be performed, as well as four actions that would act as distractor items (which, as opposed to target items, would not be performed but would be necessary for the interview phase of the study). In order to create the action lists, four lists of ten items were first created each including a combination of verbal actions, visual actions, and actions with props. Next, the actions on each list were randomly assigned as either a target action or a distractor item, using Research Randomizer[5] (a tool available to researchers interested in conducting random assignment). Of the six actions that were assigned as target items, it was ensured that 1 or 2 were actions with props, and 4 or 5 were a combination of both verbal and visual actions. There were no obvious categorical relationships between actions of any kind, and actions were all of similar developmental level. For example, any vocabulary used was comprehensible by all age groups. See Appendix B.

The clips in each segment were displayed for an average of 6.82 seconds, one after another, with a 2 second pause between clips. The order in which actors appeared in each condition, as well as the actual list of six actions that they performed, was counterbalanced in order to eliminate possible order or carryover effects. In addition, the order in which participants could be exposed to the experimental conditions (one differentiating cue/five differentiating cues) was also counterbalanced. The result of such counterbalancing procedures was a set of four videos that participants could possibly be assigned to, and were done so in a completely random order. The only aspects that were not counterbalanced were the set of actors and the lists of actions that were performed in each condition; specifically, although counterbalanced amongst themselves, the same two actors and the same two lists of actions were always associated with the same condition (e.g., one differentiating cue/five differentiating cues), so as to create a manageable

[5] www.randomizer.org

number of possible scenarios that participants could be assigned to. See Appendix C.

Distractor Task. The purpose of the distractor task was to direct participants' focus elsewhere and prevent clear memorization of the content before being presented with the main task of the experiment (i.e., the source monitoring interview). Specifically, child participants were given crayons/markers and paper, and asked to draw a picture. Each child was given approximately ten minutes to do so, and was encouraged to add to their drawing should their original product not span the desired time. Adult participants were asked to read a paragraph with content and illustrations of entirely unrelated subject matter. For adult distractor task, see Appendix D.

Source-Monitoring Interview. Participants were asked a set of pre-determined questions pertaining to the clips/target actions contained within each condition.

The first phase of the interview was in relation to the first condition that the particular participant was exposed to. Specifically, recognition questions were asked about the twelve actions performed by each of the two actors. For example, "did someone say 'I love dogs'?" The participant had the opportunity to answer *yes* or *no* the recognition questions, and responses were recorded. As a manipulation check, and in order to ensure that participants had in fact remembered the contents of each condition, recognition questions were also asked about six non-present actions (distractor items) that were not performed by the actors. Questions were asked in a completely randomized order. Moving forward, if a participant responded *no* to a recognition question, (regardless of whether it was a target action or a distractor item) the researcher moved on to the next action question. If the participant responded *yes* to a recognition question, (regardless of whether it was a target action or a distractor item) the researcher asked the participant to make a source judgment. For example, "was it Candice or Alexia?".[6,7]

The second phase of the interview was identical in all ways, however it related to the second condition that the particular participant was exposed to.

The third phase of the interview consisted of a qualitative question, reminding the participant that they mentioned remembering the actors performing some actions in the videos, and asking them to make a subjective judgment of which they found to be more difficult: remembering what the actors in the one-cue condition did, or remembering what the actors in the five-cues condition did, and why. Specifically, it asked: *Which was harder… remembering what Abby and Paige did, or what Candice and Alexia did?*

Due to the fact that there were four possible forms of the video that the participant could be randomly assigned to, there were, in turn, four versions of the interview that could be administered accordingly; one to correspond to each of the video forms. It was imperative that the researcher administer the interview that correctly corresponded with the version of the video that the particular participant watched. See Appendix E.

Procedure

The study design is outlined in Figure 1. Participation in the study took place at two elementary schools within the Waterloo Region District School Board, as well as within the Psychology department at Wilfrid Laurier University, in Waterloo Ontario.

[6] It was necessary that participants make source judgment in response to *all* recognition questions that they responded positively to (even if it was in regards to a distractor item), as their answers may shed valuable light upon the cause of their confusion or the nature of their errors.

[7] Although actual names were used when conducting the study, pseudo names were substituted for each of the actors within this document for the purpose of anonymity.

Participants were escorted to the study location by a researcher, welcomed, and thanked for their participation in the study. Once settled in, all participants were provided with an overview of the study. The overview was delivered in written form for adult participants and in oral form for child participants. Next, adult participants were asked to read and sign the letter of informed consent (which included the request for demographic information), and child participants (whose parents had already given informed consent) provided the experimenter with verbal assent (see Appendix F).

The experimental setting was equipped with a chair, a desk, and a portable computer, with no potential distractions (as best as possible). Simple headshots of each of the four actors (displaying them in a white lab coat with a neutral background) were placed on the table in front of the participant (see Appendix G). By means of a pre-determined script, the experimenter informed the participant that they would be watching some people do some things and say some things on the computer, and informed them of the actors' names. Once the participant indicated that the instructions were understood, the experimenter removed the headshots from the table and commenced the video.

Each participant was randomly assigned to one source monitoring video to watch, as described above. At no time did the researcher interrupt the participant, except to confirm when the first segment (condition) had ended and the second segment (condition) was about to begin. Once the video was complete, the experimenter played the video again in order to ensure that the participant encoded all of the target actions. When the video was complete for the second time, the participant was presented with the distractor task.

Following the distractor task, the researcher began the interview phase of the study. The researcher asked the participant if they remembered watching the people do some things and say some things on the computer. Providing that the participant agreed, the researcher returned the headshots to the table and asked the participant if they could remember the actors' names. Again, the headshots did not display the actors in character, so as to ensure that participants were not reminded of any of the visual cues that they exhibited during the video. If the participant correctly identified each actor by name, the researcher moved on. If the participant identified any of the actors incorrectly, the researcher was sure to remind the participant of their names. Finally, once the participant could match a name to each of the actors' faces, the researcher let the participant know that they had a list of all of the things that occurred in the video, and wanted to see if they could remember who did them.

The source-monitoring interview always began with the first condition that the participant was exposed to in their video. The researcher instructed the participant to think back to that specific part of the video, and removed the headshots of the actors that were not contained within the condition from the table in front of them. The researcher asked all of the questions specified on the interview sheet for the particular video that the participant watched (including both recognition and source-monitoring questions for target actions and distractor items). Once the researcher had completed questions pertaining to the first condition, they did the same for the second condition, and administered the final qualitative question. Upon completion of the interview, the researcher answered any remaining questions that the participant may have had, and thanked them for their participation in the study.

Coding

See Appendix H. All data were double coded by two Research Assistants in the Child Memory Lab at Wilfrid Laurier University. Coders were trained by the Principal Investigator, and any disagreements were handled by re-coding the entire interview in question until agreements were reached about the discrepancies. Inter-rater reliability was calculated at one hundred percent (inclusive of all study interviews).

Recognition Scores/Recognition Coding. Coding of each interview was broken down by condition. For each condition that the participant was exposed to, coders first noted the overall number of target item recognition questions that they answered correctly. In other words, they noted the number of instances that the participant remembered an action that was actually present in the segment (for example, answering "yes" to the question 'did someone say "I love dogs"?'; see Appendix B). The purpose of this was to ensure that participants had encoded the actions presented within the condition. As the actions that appeared within each condition were based on two action lists (each containing 6 target actions and 4 distractor items), the maximum recognition score was 12. For the purpose of further analysis into potential responses biases, the number of correct target item recognition questions was also broken down by source in the particular condition. For example, for the five-cues condition, coders noted the total number of target item recognition questions that the participant answered correctly when the source who performed the particular action was Candice (out of 6), as well as the total number of target item recognition questions that they answered correctly when the source who performed the particular action was Alexia (out of 6).

Source Scores/Source Coding. Next, coders recorded the number of correct source attributions out of the overall number of target item recognition questions correct. In other words, they recorded the number of instances when the participant correctly identified the source of actions they remembered being performed in the segment (e.g., after acknowledging that they remembered someone saying "I love dogs", whether the participant correctly identified if it was Candice or Alexia). This was to assess the participant's memory for source, separate from recognition memory ability. Again, for the purpose of further analysis into potential response biases, the number of correct source attributions was also broken down by source. For example, for the five-cues condition, coders recorded number of instances that the participant correctly identified Candice as the source, as well as the number of instances that the participant correctly identified Alexia as the source.

Misleading Recognition (Distractor) Scores/Coding. Finally, coders noted the number of distractor items that were correctly identified. In other words, they noted the number of instances that the participant correctly identified an action that was not present as being just that— not present within the segment. For example, answering "no" to the question "did someone blow up a balloon?" (see Appendix B). The purpose of this was to incorporate a memory manipulation check, and to further assess the participants' memories for actions presented within the conditions. As the actions in each condition were based on two action lists (each containing 6 target actions and 4 distractor items), the maximum score was 8. For the purpose of further analysis into potential response biases, the distractor items that were incorrectly identified were broken down by source. For example, for the five-cue condition, coders noted the number of instances that Candice was identified as the source of a distractor, as well as the number of instances that Alexia was identified as the source of a distractor; though neither of them actually performed these actions. See Appendix G.

All data, including the final qualitative interview question (asking participants to make a subjective judgment of which they found to be more difficult: remembering what one pair of actors did [the actors in the one-cue condition] or what the other pair of actors did [the actors in the five-cues condition], and why) were entered into SPSS for statistical analyses.

Results

Analytic Strategy

First, preliminary analyses were conducted to search for unintended differences. Next, recognition analyses were conducted to determine participants' memory for event details, separate from source monitoring performance. The foremost and central portion of the analysis, the inferential analysis, was then conducted to explore the explicit hypotheses of the study which revolved around age differences in source monitoring ability. Specifically, analyses were conducted to determine whether the present sample displayed an overall developmental progression in source monitoring ability, as well as to determine whether further age differences were present when the number of visual cues that differed between two sources was manipulated; one-cue and five-cues. Finally, the qualitative data regarding subjective difficulty (and purported reasons for such difficulty) were addressed.

For the purpose of data analysis, proportion scores were computed for each participant (i.e. proportion of source monitoring score divided by recognition memory score) for each condition to create a score that was a true reflection of source monitoring ability, separate from recognition memory ability. These proportion scores acted as the main measure when referring to source monitoring scores for each of the conditions.

Preliminary Analysis/Data Check

Primacy/Recency Analysis. The purpose of a primacy/recency analysis is to determine whether the serial position of a condition had any significant effect on a specific outcome or measure. Although counterbalancing and random assignment methods were used, and personal characteristics should statistically even out across conditions, a preliminary analysis was conducted to determine whether there were any primary or recency effects of cue condition. Specifically, one analysis was conducted to search for potential effects of each condition's presentation order on participants' score in that condition, and another was conducted to search for potential effects of each condition's presentation order on participants' scores in each of the two conditions.

Effect of condition presentation order on relative condition score. Six independent samples *t*-tests were run to search for potential effects of each condition's presentation order on participants' score in that condition. The first three tests (one for each age group) were run to search for significant differences between scores in the *one-cue condition* for those who were exposed to the one-cue condition first (and the five-cues condition second), and those who were exposed to the one-cue condition second (and the five-cues condition first). Higher scores for those who were exposed to the one-cue condition first would indicate a primacy effect for the one-cue condition, and higher scores for those who were exposed to the one-cue condition second would indicate a recency effect of the one-cue condition. Results indicated that there were no significant differences for the 3-5-year-old or the 18-21-year-old age groups; $t(28) = 1.01$, $p = .32$ and $t(30) = .46$, $p = .65$ respectively. That being said, a significant difference was found to exist for the 6-8-year-old-age group, as those who were exposed to the one-cue condition first had a mean score in the one-cue condition ($M = .84$, $SD = .11$) that was significantly higher than those who were exposed to the five-cue condition first ($M = .64$, $SD = .15$), $t(35) = 4.77$, $p = <.001$.

The next 3 tests (one for each age group) were run to search for significant differences between scores in the *five-cues condition* for those who were exposed to the five-cues condition first (and the one-cue condition second), and those who were exposed to the five-cues condition second (and the one-cue condition first). Higher scores for those who were exposed to the five-cues condition first would indicate a primacy effect of the five-cues condition, and higher scores for those who were exposed to the five-cues condition second would indicate a recency effect of the five-cues condition. Results indicated that there were no significant differences for any age group; 3-5-year-olds $t(28) = -.56$, $p = .58$, 6-8-year-olds $t(35) = -1.67$, $p = .10$, 18-21-year-olds $t(30) = .67$, $p = .51$. Accordingly, there were no primacy or recency effects of condition presentation order on source monitoring scores in the five-cues condition.

Thus, one out of the 6 *t*-tests showed a statistically significant effect (the 6-year-olds in the one-cue first condition). Seeing as counterbalancing procedures were used, no primacy or recency effects were predicted, and only approximately half of the participants in the 6-8-year-old age group were exposed to the one-cue condition first, this isolated effect did not raise concern within the scope of the study.

Effect of condition presentation order on between-condition scores. A second set of analyses was run to test for potential effects of condition presentation order on *between-condition* scores. Specifically, a 2(cue condition: one-cue, five-cues) x 3(age: 3-5, 6-8, 18-21) x 2(condition position: one-cue presented first, five-cues presented first) ANOVA was run, and a significant three-way interaction was found to exist between cue condition, age, and condition position, $F(2, 93) = 5.93$, $p = .004$, , $\eta^2_p = .11$. As such, six post hoc paired-samples *t*-tests were conducted to investigate this significant interaction, one for each age group comparing scores between the two cue conditions when they were exposed to the one-cue condition first, and another for each age group comparing scores in each of the cue conditions when they were exposed to the five-cues condition first. The only significant differences between scores in the two cue conditions that were found to exist were for the 3-5-year-olds and the 6-8-year-olds who were exposed to the five-cues condition first; $t(13) = -3.01$, $p = .01$, and $t(17) = -5.13$, $p < .001$ respectively. These results indicated that both the 3-5-year-olds and the 6-8-year-olds who were exposed to the five-cues condition first had significantly higher source monitoring scores in the five-cues condition (3-5 $M = .76$, $SD = .17$; 6-8 $M = .86$, $SD = .14$) than in the one-cue condition (3-5 $M = .65$, $SD = .24$; 6-8 $M = .64$, $SD = .15$).

Thus, two out of the 6 *t*-tests showed a statistically significant effect (the 3-5-year-olds and the 6-8-year-olds who were exposed to the five-cues condition first had significantly higher source monitoring scores in the five-cues condition). As noted in the first primacy/recency analysis, as counterbalancing procedures were used, no primacy or recency effects were predicted, and only approximately half of the participants in the 6-8-

year-old age group were exposed to the one-cue condition first, the effect did not raise concern within the scope of the study.

Gender Analysis. In order to determine whether any gender differences existed in overall source monitoring performance, data were broken down by gender and a 3(age: 3-5, 6-8, 18-21 years) x 2(gender: female, male) analysis of variance (ANOVA) was conducted with overall source monitoring proportion as the dependent variable. A main effect of gender was found to exist among participants, $F(1,93) = 8.15$, $p = .005$, $\eta^2_p = .08$. Specifically, females scored slightly higher than males in all age groups. See Table 4. As gender was not considered a factor within the scope of the study and past research has not been known to find such an effect, gender differences were likely a product of the specific sample and were not included as a factor in subsequent analyses.

Recognition Analysis

Recognition of Target Actions. An analysis was conducted to investigate how many of the target actions participants actually remembered occurring within each condition. This was important to analyze as these were the target actions that source-monitoring judgments were made in regards to for each condition, as well as one of the factors that was used to compute the proportion score used within the study as the main measure of source monitoring ability. In terms of the one-cue condition, out of a possible total of 12, 33% of individuals in the 3-5-year-old age group recalled all 12 target actions having occurred, 76% recalled at least three-quarters, and their overall mean score of 9.73 ($SD = 2.65$) was significantly greater than chance. For statistics regarding chance, see Table 5. Furthermore, 35% of 6-8-year-olds recalled all target actions having occurred, 95% recalled at least three-quarters, and their overall mean score of 10.70 (*SD* = 1.56) was significantly greater than chance. Finally, 34% of 18-21-year-olds recalled all target actions having occurred, 97% recalled at least three-quarters, and their overall mean score of 10.88 (*SD* = 1.18) was significantly greater than chance. In terms of the five-cues condition, similarly out of a possible total of 12, 20% of individuals in the 3-5-year-old age group recalled all 12 target actions having occurred, 80% recalled at least three-quarters, and their overall mean score of 10.03 (*SD* = 1.87) was significantly greater than chance. Furthermore, 30% of 6-8-year-olds recalled all target actions having occurred, 84% recalled at least three-quarters, and their overall mean score of 10.35 (*SD* = 1.74) was significantly greater than chance. Finally, 34% of 18-21-year-olds recalled all target actions having occurred, 97% recalled at least three-quarters, and their overall mean score of 10.56 (*SD* = 1.52) was significantly greater than chance.

In order to determine whether there were any major statistical discrepancies between recognition scores in either of the conditions or between any age groups, a 3(age: 3-5, 6-8, 18-21 years) x 2(cue condition: one-cue, five-cues) ANOVA with a mixed factorial design was run. Age was the between-subjects variable, cue condition was the within-subjects variable, and target action recognition score was the dependent variable. The interaction was not significant, indicating that target action recognition scores were not significantly higher or lower in one-cue condition or the other as a function of age group, $F(2,96) = .91$, $p = .41$. Furthermore, there was no main effect of age group, $F(2,96) = 2.89$, $p = .06$, indicating that participants in each of the age groups did not differ significantly in their overall ability to recognize target actions, and there was no main effect of cue condition, $F(1,96) = .31$, $p = .58$, indicating that (overall) participants were able to recognize the same number of target actions in each of the cue conditions. Thus, any developmental differences in source monitoring cannot be due simply to recognition effects.

Recognition of Distractor Items. An analysis was conducted to investigate how many of the distractor items participants correctly rejected (i.e., correctly identified as not being present) within each condition. This was not simply the opposite of the above recognition of target actions analysis, but instead an important measure that gave an indication of participants' ability to discriminate between present and non-present items, as well as illuminate any potential response biases. Regarding the one-cue condition, out of a possible total of 8, 50% of individuals in the 3-5-year-old age group correctly identified all 8 distractor items, 80% identified at least three-quarters, and their overall mean score of 6.63 (*SD* = 2.03) was significantly greater than chance. For statistics regarding chance, see Table 6. Furthermore, 70% of 6-8-year- correctly identified all 8 distractor items, 97% identified at least three-quarters, and their overall mean score of 7.59 (*SD* = .73) was significantly greater than chance. Finally, 94% of 18-21-year-olds correctly identified all 8 distractor items, 100% identified at least three-quarters, and their overall mean score of 7.91 (*SD* = .39) was significantly greater than chance. Regarding the five-cues condition, similarly out of a possible total of 8, 30% of individuals in the 3-5-year-old age group correctly identified all 8 distractor items, 80% recalled at least three-quarters, and their overall mean score of 6.43 (*SD* = 1.70) was significantly greater than chance. Furthermore, 60% of 6-8-year-olds correctly identified all 8 distractor items, 95% recalled at least three-quarters, and their overall mean score of 7.35 (*SD* = 1.06) was significantly greater than chance. Finally, 75% of 18-21-year-olds correctly identified all 8 distractor items, 100% recalled at least three-quarters, and their overall mean score of 7.69 (*SD* = .59) was significantly greater than chance.

In order to determine whether there were any major statistical discrepancies between recognition of distractor items in either of the conditions or between any age groups, a 3(age: 3-5, 6-8, 18-21 years) x 2(cue condition: one-cue, five-cues) ANOVA with a mixed factorial design was run. Age was the between-subjects variable, cue condition was the within-subjects variable, and distractor item scores was the dependent variable. The interaction was not significant, indicating that distractor item recognition scores were not significantly higher or lower in one-cue condition or the other as a function of age group, $F(2,96) = .02$, $p = .98$. That being said, there was a main effect of age group, $F(2,96) = 10.58$, $p < .001$, indicating that there was a significant difference between age groups in overall ability to reject distractor items. Post hoc tests of multiple comparisons went on to show that both the 18-21-year-olds (*M* = 7.80, *SE* = .20) and the 6-8-year-olds (*M* = 7.47 , *SE* = .18) correctly rejected more distractors than the 3-5-year-olds (*M* = 6.53 , *SE* = .21), although no difference existed between the 18-21-year-olds and the 6-8-year-olds, Scheffe < .05. There was also a main effect of cue condition, indicating that (overall) participants were able to reject distractor items significantly more often in the one-cue condition (*M* = 7.38, *SE* = .12) than in the five-cue condition (*M* = 7.16, *SE* = .12), $F(1,96) = 6.46$, $p = .01$.

Although there was a significant main effect of age group, this main effect follows a normal developmental pattern, and thus was not a concern for researchers. The main effect of cue condition, although interesting, was also not a concern, as all participants performed considerably well regardless.

Main Analysis

The main analysis of the study targeted source-monitoring ability. The statistical test used to explore hypotheses was a 3(age: 3-5, 6-8, 18-21 years) x 2(cue condition: one-cue, five-cues) ANOVA with a mixed factorial design, with age as the between-subjects variable, cue condition as the within-subjects variable, and source monitoring proportion as the dependent variable.

Hypothesis 1. Hypothesis 1 (whether there was an overall/general developmental progression in ability to make accurate source judgments) was investigated by determining whether or not the ANOVA yielded a significant main effect of age. With a *p*-value of < .001 there was strong evidence against the null, and it could be concluded that there was a main effect of age group (at least one of the groups differed significantly from another), $F(1,96) = 15.91$, $p < .001$, $\eta^2_p = .25$. In light of this significant main effect, post hoc tests of multiple comparisons were conducted to determine the direction and strength of the progression. Significant mean differences were found to exist, with the 18-21-year-olds ($M = .90$, $SE = .02$) scoring significantly higher than both the 6-8-year-olds ($M = .78$, $SE = .02$) and the 3-5-year-olds ($M = .72$, $SE = .02$), Scheffe $p = .001$ and Scheffe $p < .001$ respectively. In terms of the younger age groups, although the 6-8-year-olds performed slightly better overall than the 3-5-year-olds, the difference did not reach significance, Scheffe $p = .16$. Overall, in line with previous research, there was sufficient evidence to conclude that there was an overall/general developmental progression in participants' ability to make accurate source judgements, and hypothesis 1 was accepted. See Figure 2.

Hypothesis 2. In order to determine whether adults were more accurate in the five-cue condition than in the one-cue condition, whether 3-5-year-olds found a single cue more salient/distinctive than a greater number of differences, and whether 6-8-year-olds' scores differed in the one- and five-cue conditions, the above mentioned 3(age: 3-5, 6-8, 18-21 years) x 2(cue condition: one-cue, five-cues) ANOVA was examined further.

The ANOVA was tested for a significant interaction, which would indicate whether the pattern of the scores in each of the cue conditions differed as a function of age group, as predicted. The interaction was not significant, $F(2,96) = .13$, $p = .88$. Instead, results revealed a significant main effect of cue condition, $F(1,96) = 7.76$, $p = .006$, $\eta^2_p = .08$. Specifically, overall, scores in the five-cues condition ($M = .82$, $SE = .02$) had a mean that was significantly higher than scores in the one-cue condition ($M = .77$, $SE = .02$).

In light of the significant main effect of cue condition, there was sufficient evidence to conclude that (on average) 18-21-year-olds make significantly more accurate source monitoring judgments when there are more distinguishing visual source cues/when sources are less visually similar (i.e. more accurate in the five-cue than in the one-cue) as predicted by the hypothesis. See Table 7. Although scores in both conditions were close to ceiling, an effect size was calculated in order to facilitate interpretation of the significance and determine the strength of the difference between scores in the two conditions, $d = .42$. This is a medium effect size, according to Cohen's (1988) guidelines.

As the above mentioned 3(age: 3-5, 6-8, 18-21 years) x 2(cue condition: one-cue, five-cues) ANOVA yielded a significant main effect of condition (with greater mean scores in the five-cues condition) as opposed to a significant interaction, it is clear that the 3-5-year-olds did not earn scores in the one-cue condition that were higher than those in the five-cues condition as the hypothesis had originally predicted. As a result, there was no evidence to conclude that 3-5-year-olds make significantly more accurate source monitoring judgments when there are less distinguishing visual source cues contained within the event/when sources are more visually similar but contain a distinctive visual cue as was predicted. See Table 7.

As the above mentioned 3(age: 3-5, 6-8, 18-21 years) x 2(cue condition: one-cue, five-cues) ANOVA yielded a significant main effect of cue condition, it is clear that the 6-8-year-olds earned scores in the five-cues condition that were higher than those in the one-cue condition. As a result, there was sufficient evidence to conclude that (on average) 6-8-year-olds make significantly more accurate source monitoring judgments when there are more distinguishing visual source cues/when sources are less visually similar. See Table 7.

Subjective Difficulty Analysis

An analysis of subjective difficulty was run to determine which pair of sources participants in each age group *alleged* to be harder to remember: sources in the one-cue condition, or sources in the five-cues condition. Results of the frequency analysis revealed that 50 percent of participants in the 3-5 year-old age group found the pair of sources in the one-cue condition harder to remember, as opposed to 40 percent who found the sources in the five-cues condition harder and ten percent who replied *both/neither*[8]. These subjective results were consistent with their performance, as participants performed slightly better in the five-cue condition than the one-cue condition. As for the 6-8-year-old age group, results revealed that 41 percent of participants found the pair of sources in the one-cue condition harder to remember, as opposed to 57 percent who found the sources in the five-cues condition harder (three percent replied *both/neither*). These subjective results did not parallel the age group's performance, as participants performed considerably better in the five-cue condition than the one-cue condition. Finally, results revealed that 59 percent of participants in the 18-21-year-old age group found the pair of sources in the one-cue condition harder to remember, as opposed to 34 percent who found those in the five-cues condition harder (six percent replied *both/neither*). These subjective results were in line with the age group's performance, as participants performed significantly better in the five-cue condition than the one-cue condition.

Perceived Reasons for Difficulty. In order to further investigate the subjective difficulty of one condition over another for each age group, answers to the qualitative question "*who was harder to remember [sources in the one-cue condition or sources in the five-cues condition]?"* were analyzed and the most common answers were documented. As for the 3-5-year-olds, 50 percent of whom reported sources in the one-cue condition to be harder to remember, the most commonly noted reasons for difficulty included: *had to think hard*, *couldn't remember/harder to remember*, *because of (specific actions)*, *I don't know*, and *they did harder things*. The pattern of what they *believed* to be more difficult did in fact line up with how they actually performed; they did not

[8] Both/neither was not an actual response option; it was only recorded as the participants' response if they insisted that they were not able to make a decision between the two pairs of sources.

perform quite as well in the one-cue condition (which they believed to be more difficult) than they did in the five-cues condition. As for the 6-8-year-olds, 57 percent of whom believed sources in the five-cues condition to be harder to remember, the most commonly noted reasons for difficulty included: *Abby and Paige were easier*, *I don't know,* references to differences in clothing (e.g., hat/necklace), and references to order of presentation (e.g., they came first/second). The pattern of what they believed to be more difficult did not line up with how they actually performed; instead, as noted above, they were actually more accurate in the five-cues condition (which they believed to be more difficult) then they did in the one-cue condition. Finally, as for the 18-21-year-olds, 59 percent of whom reported sources in the one-cue condition to be harder to remember, the most commonly noted reasons for difficulty included: references to order of presentation, references to similarity/dissimilarity of clothing and features, the fact that dissimilar sources [in the five-cues condition] were easier to differentiate/more memorable. The pattern of what they believed to be more difficult did in fact line up with how they actually performed; they did not perform as well in the one-cue condition (which they believed to be more difficult) as the five-cues condition.

Discussion

The purpose of the present study was to evaluate age differences in source monitoring ability, as well as to look further into the issue of source similarity and determine whether children benefit more from a single distinctive visual cue than a greater number of cues; essentially, to better understand the developmental differences in the effect of cues on children's source-monitoring performance. The findings with regards to these questions are discussed first, followed by a consideration of secondary questions that were addressed by the data. Practical implications, limitations, and future research are discussed.

Age Differences. One of the main hypotheses of the study was that there would be an overall/general developmental progression in participants' ability to make accurate source judgments. Specifically, as children grow older their ability to accurately monitor source should increase, with marked improvements likely to occur in the 3-8-year-old age range. The rationale for this hypothesis was based on the highly documented and well-replicated finding in the literature that source monitoring ability improves with age (Roberts, 2002), and was incorporated as a hypothesis into the present study to ensure that the present sample displayed the same developmental trend as samples in the past. Results of the statistical test indicated that the present sample did indeed display an overall/general developmental progression in participants' ability to make accurate source judgements. This did not speak to participants' performance in one condition or another but, rather, their ability to monitor source in general. As a result of the observed overall/general developmental progression, hypothesis one was accepted.

Acceptance of hypothesis one suggests two important implications. First, it adds a valuable contribution to the current body of source monitoring literature by offering yet another example of a that study has replicated similar results to those that have been found numerous times by numerous researchers. The more a particular result is replicated, the more assured we can be about its authenticity and, thus, the more we will know about the corresponding construct as a whole. Second, acceptance of this hypothesis allowed for assurance that there were no major inconsistencies with the present sample compared to those of source monitoring studies in the past, and that further analyses based on it were warranted and acceptable. What should be noted when discussing the present sample, however, was that although the results did illustrate a developmental trend (with adult participants performing significantly better than child participants in both the 3-5- and 6-8-year-old age ranges), the difference between average scores of children in the 3-5- and 6-8-year-old age range did not reach significance. This result was surprising due to the fact that marked improvement was expected to occur within the 3-8-year-old age range (Roberts, 2002). However, as 6-8-year-olds still performed *slightly* better than the 3-5-year-olds, it was not seen as cause for concern and was anticipated to be a simple result of sample size and individual difference characteristics.

When discussing age differences, one important point that ought to be highlighted, although separate from source-monitoring ability, is how well the 3-5-year-old participants performed in the overall recognition task. Although the literature suggests that young children may not be as efficient as older children or adults in the execution of memory tasks in general, analysis of their performance in regards to recognition of target actions revealed that they performed very well (significantly greater than chance in each cue condition) and quite similarly to their older child and adult counterparts. It is possible that the adult participants may have been performing close to ceiling and, therefore, that the 3-5-year-olds' scores may not be as similar on a larger scale. Nonetheless, it is clear that young children may be better in the execution of some memory tasks than may otherwise be expected. In terms of the wider framework, this is important information as it highlights the abilities of young children and their competence in terms of recognition memory at a short delay.

Condition Differences. The second major hypothesis of the study was that there would be developmental differences in the effect of cues on children's source monitoring performance. Specifically, according to the similarity effect, adults would be more accurate in the five-cue condition than in the one-cue condition because the larger number of differences between sources decreases their similarity. Therefore, the easier it is to distinguish between the sources. Regarding the 3-5-year-olds, source scores may not follow the same pattern predicted by the similarity effect. Due to young children's limited cognitive capacity, even though there are a greater number of differences in the five-cue condition than the one-cue condition, preschoolers may not have the cognitive development factors (e.g., working memory, inhibitory control) necessary to use multiple cues effectively. As such, they may find a single visual cue more distinctive than a greater number of differences. Data from the 6-8-year-olds was exploratory as it was unclear whether and how their scores would differ in the one- and five-cue conditions. Therefore, no specific prediction was made.

In terms of the adult participants, higher mean scores were recorded and source monitoring judgments were more accurate when more distinguishing visual source cues existed between sources/when sources were less visually similar (in the five-cue condition). The rationale for this hypothesis was based on the assumptions of the similarity effect, as well as past research that has demonstrated adults to be significantly superior at source monitoring judgments when more differences exist between sources (Roberts, 2002). As a result of the higher scores in the five-cue condition, this hypothesis was accepted. Similar to the motivation for confirming normative trends in age differences for overall performance, it was important to confirm that the adult portion of the present sample adhered to historical trends in source monitoring ability as well in order to indicate that there were no major inconsistencies with the present sample compared to those of source monitoring studies in the past, as well as to have an appropriate comparison group upon which to determine and define developmental trends within the child samples.

In terms of the 3-5-year-old participants, although the rationale for the current study predicted a single visual cue to be more distinctive than a greater number of differences (i.e., the *opposite* pattern than that predicted by the similarity effect for adults), slightly higher mean source monitoring scores were recorded in the five-cues condition than the one-cue condition. Similar to the adult participants, individuals in this age group appeared to benefit more from multiple differences between sources as predicted by the similarity effect as opposed to a single visual cue. It is important to note, however, that although young children performed significantly better in the five-cues condition, they performed relatively well (above chance) in both of the cue conditions.

Due to the fact that the obtained outcome was not the expected outcome for this age group, and knowing that both age groups performed better in the five-cues condition than the one-cue condition, researchers went on to investigate further by determining the magnitude of the difference between mean scores in each of the two conditions. The purpose of doing so was to determine whether there was *as* large a difference between scores in the two conditions for the 3-5-year-olds as there was for the adult participants. If there was not, it may be evidence that 3-5-year-olds do not benefit *as much* from multiple cues as adults do, the proposed theory may still be possible, and evidence for a developmental trend within the results may still exist. As anticipated, careful inspection of effect sizes indicated that the magnitude of the difference between scores in the one-cue and five-cues condition for the 3-5-year-olds (*d* = .31) was much smaller than for the adult participants (*d* = .41). Although such effect sizes are not typically calculated in light of the non-significant interaction, they were included in the present discussion as they display a tendency that is somewhat supportive of the proposed theory. In other words, they are in the direction of what was expected, and encourage follow-up.

In terms of the 6-8-year-old participants, in which no specific prediction about performance was made, higher mean scores were recorded in the five-cues condition than the one-cue condition. As such, participants in this age group benefited more from multiple source cues than they did one distinctive visual cue.

Summary. Overall, each age group demonstrated higher mean scores in the five-cues condition than the one-cue condition. Taken together, the most evident message that can be taken from these results is that multiple cues, in the context of the current study, may be more beneficial to individuals than a single visual cue (as predicted by the similarity effect). Although this conclusion is in line with our prediction for adults, it is contrary to that predicted by the proposed theory for young children.

As the present study was exploratory in nature, further research and replication may be necessary to determine the exact effect of single distinctive versus multiple visual cues on children's source monitoring performance. Although it was not the case in the context of the present study, the researchers continue to propose that it may be possible for the number and type of cues in a source monitoring simulation to be manipulated in such a way that the 'similarity effect' (that has so often been found to exist in adult populations and supposed to exist for children) may not hold true, or do so to the extent that may once have been anticipated. In other words, *the more* cues may not always be *the better*. As noted by Lindsay and colleagues (1991), children may attend to different aspects of events, and it may be that the kinds of memory records that quickly and easily come to mind when remembering an event differ for children and adults. In addition, it may be possible is that what creates an accessible memory record for individuals to use when making a source-monitoring decision may be different for children than it is for adults and, as source-monitoring errors are much more likely to degrade when memories are vague, it is extremely important to find out what children find distinct or memorable in order to further understand the nature of their errors as well as to understand developmental differences.

There are various reasons for which the results of the present study may have been null and for which the proposed theory, which suggests that a single cue may be more beneficial for young children than a greater number of cues, may still be possible. First and foremost, it is possible that *t-shirt colour* is not a single cue that presents as distinct enough to aid children (more so than multiple cues) in a source-monitoring situation. Although children did perform above chance when using it as a single cue, for one reason or another it may not be a cue that is effectively encoded, that makes its may into working memory, or that becomes an accessible memory record upon which to use when making a source-decision, more so than multiple cues. Second, it is possible that the type of single cue in general (*visual*) did not present as distinct enough to work on its own and support source monitoring in young children. Perhaps, such as in the case of Lindsay and colleagues' auditory cue (1991, Experiment 1), single cues perceived through other modalities are more helpful than single visual cues for young children. Finally, it may be possible that what aids young children and presents as even more distinct than a single visual cue is a *single cue that represents a concept or construct*. As noted previously, the single cue that was found to aid young children with source monitoring in the Lindsay and colleagues (1991) study was *gender*. Although gender has the ability to display itself as a single cue, it has a different type of distinctiveness about it that comes along with much more knowledge, experience, familiarity, and inherent detail than *t-shirt colour* alone. There are various points in the academic and empirical literature that can be used to support the notion that gender, even when presented as a single entity, may be a special or stronger type of cue. For example, Bussey and Bandura (1999) state that human differentiation on the basis of gender is a fundamental phenomenon that affects virtually every aspect of people's daily lives, and that gender conceptions are constructed from the complex mix of experiences and how they operate in concert with motivational and self-regulatory mechanisms to guide gender-linked conduct throughout the life course. In addition, according to Bem (1983), sex differences are naturally and inevitably more perceptually salient to children than other differences, and our culture does not construct any distinctions between people that we perceive to be as compelling as sex. It is clear that this description of gender may not be consistent cross-culturally or constant across age groups, but it is certainly characteristic of young westernized children of our targeted developmental standing. As a result, it is possible that although preschoolers may not have the cognitive development factors (e.g., working memory, inhibitory control) necessary to use multiple cues effectively in source-monitoring situations, they may be able to make use of stronger or complex cues when presented as a single entity, without putting too much of a demand on their cognitive resources.

Practical Implications and Directions for Future Research

The present study sets forth various areas of practical implication and future research. First, the present study proposed the notion that the similarity effect (or, the idea that ‘the more differences between sources the better’) ought not be generalized to all individuals in all source-monitoring situations to the same extent. Discussion suggested that making such generalizations may not be appropriate due to the fact that single cues such as gender in the Lindsay et al. (1991, Experiment 1) study exist that may be more beneficial to young children than multiple cues for source monitoring. The present study proposed this as a theory, and attempted to start discussion in this area by testing a single cue (based on its visual qualities) for potential distinctiveness to children. Although this cue was not found to be as advantageous as anticipated, future research should be directed toward determining other cues or constructs that young children find distinct. Determining such cues would not only highlight the abilities of young children and speak to their ability to monitor source to a similar extent as adults, but would increase their credibility in critical situations such as forensic or eyewitness accounts. Most practically, it could aid other researchers and professionals in the area of source-monitoring training by highlighting information that is particularly important for young children to consider when recalling events, and that should be attended to when responding to questions about events witnessed from different sources (Thierry & Spence, 2002).

Second, a great deal of the foundation of the proposed theory was based on the results of Lindsay and colleagues' (1991, Experiment 1) study, which found one distinct difference between sources to be extremely beneficial for young children. That being said, Lindsay and colleagues had not tested the effect gender as a single cue against the effect of multiple cues. Instead, they had tested it against a more similar condition (two female voices, as opposed to a male and a female voice) and concluded that voices of different genders were more beneficial than voices of the same gender (in essence, in support of the similarity effect). Although they may have interpreted these results in such a way, and gender may have been the *more different* condition in that case, what the researchers of the present study find more important was the fact that gender was a *single cue* and was *distinct enough to benefit young children* in a source-monitoring situation. From this landmark study it is clear that it is not necessarily the number of cues that is important for young children, but rather the distinctiveness of the cues—which certainly makes sense, considering the cognitive capabilities of young children. Future research should be directed toward testing gender specifically as a single difference against multiple differences between sources, in order to concretely and formally establish it as a distinctive cue for young children.

Third, although the method of the present study represented the most practical way to obtain an accurate indication of source monitoring ability separate from memory ability (i.e., recognition question followed by a source question for instances when a memory was recalled), it did not necessarily grant the researchers a qualitative perspective into the mental processes that any particular participant may have been using to make a source judgment, including the way that the decision making process may have played-out in their mind, any inner-dialogue that may have occurred, the decision criteria that they may have used/the automatic processes that they may have succumbed to, or the potential self-awareness or introspective knowledge of their own memory capabilities (meta memory). Future research in this area should incorporate a *think aloud procedure* in which participants are encouraged to vocalize their experience of the decision-making process. This would be an excellent measure to use in order to further explore developmental differences and understand more about children's source monitoring abilities, failures, and the processes they engage in.

Along the same line, future research should be directed toward a more in-depth qualitative analysis of participants' subjective difficulty. Although a qualitative question was incorporated into the present study and participants were asked which pair of actors they found to be more difficult and why, subsequent studies should look further into this area and gain a more in-depth understanding of their reasons for answering in such a way. For example, in the present study, child participants were often unable to articulate why they found one pair of sources to be more difficult to differentiate than the other. Adult participants, on the other hand, were able to provide more sophisticated answers including references to order of the cue conditions as well as references to similarity/dissimilarity of clothing and features and the fact that dissimilar sources were easier to differentiate/more memorable (see Appendix I). With the appropriate interview tools and techniques incorporated into future studies, researchers could gain a much more rich understanding of what is happening in the minds of participants when making source decisions in such a context.

Although counterbalancing and random assignment methods were used in the context of the current study, and personal characteristics should statistically even out across conditions, the primacy/recency analysis that was conducted as a part of the preliminary analysis yielded some interesting and important results. Specifically, in regards to the effect of condition presentation order on relative condition score (i.e., potential effects of each condition's presentation order on participants' score in that condition), a significant difference was found to exist for the 6-8-year-old-age group as those who were exposed to the one-cue condition first had a mean score in the one-cue condition that was significantly higher than those who were exposed to the five-cue condition first. In regards to the effect of condition presentation order on between-condition scores (i.e., potential effects of condition presentation order on scores that participants achieved in each of the two conditions), 3-5-year-olds and 6-8-year-olds who were exposed to the five-cues condition first had significantly higher source monitoring scores in the five-cues condition than in the one-cue condition. As aspects of the significant results highlighted by these analyses could arguably be in line with the proposed theory, future research should be directed toward further investigating and understanding such primacy/recency effects in this context.

Finally, future research should repeat the study with a greater delay as well as with a non-repeated-measures design, in order to determine the effects of such methodological changes. It may also be important to investigate the potential effects of competing pre- and post-event information, and culture.

Limitations

As with most academic or empirical literature, the present study was not without limitations. The most general limitation of the present study was in regards to the generalizability of standardized lab-based studies, and the applicability of their results to real-life settings. Although as many precautions as possible were taken in order to isolate the construct of interest and present it in a way that was clear and precise, the study circumstances may not parallel the way it may present itself in day-to-day conditions, particularly in regards to impeding influences from the environment, peers, or other mental processes. As such, although researchers may be aware of the way that results played out in the laboratory, it is important consider that they are specific to the context of the study from which they resulted, and to offer such results with caution. Along the same lines, it is important to note that all target actions and distractor items used in the study were either positive or neutral in nature, and never negative. As such, when making inferences about how the results may apply to real-world settings and situations, it is important that they are limited to those that are positive or neutral in nature, until research on negative situations is undertaken.

Another important limitation of the current study is that cognitive control variables were not used. For example, baseline readings of working memory, inhibitory control, and other such constructs that have been known to be implicated with source monitoring were not taken upon commencement of participation. As a result, although participants were sorted into appropriate age groups, there is no way of knowing whether particular participants had deficits in required areas or were of different developmental standings, potentially affecting the results. In addition, cognitive control variables may have served as important tools for discovering/replicating correlations between particular constructs and source-monitoring ability.

Although random assignment procedures were used and conditions were as counterbalanced as was practical, convenience sampling is a reality of many studies in which random sampling of the population is not feasible. The child participants in the present study were all recruited from one of two public elementary schools located in an upper-middle class suburban neighbourhood of a large town, in close proximity to one another. Although it is not clear if or how this may have affected the results, it is likely that many of the participants from this portion of the sample shared many individual difference characteristics and had more in common than they would had they been randomly selected from a population covering a wider range or demographic. In addition, the adult participants enrolled in the study were students all attending a reputable university and completing the study in exchange for course credit. Among other variables, it is clear that these individuals were on the moderate to high end of the intelligence spectrum, and may not be precisely representative to the population of 18-21-year-olds as a whole.

Conclusion

The present study theorized that the similarity effect, as it is known to exist in the literature, may not be generalizable to the same extent to all age groups across all dimensions. Although it did not yield significant results, it provided a detailed justification for why one or fewer distinct cues may be more beneficial to young children's source-monitoring than many, and initiated discussion and research in this area by suggesting the significance of gender as a cue and testing the influence of a salient visual cue. Overall, in order to help children perform to their best ability and recall memories that are vivid and complete, it is important to determine what they find distinct and least susceptible to deration over time. Studies of developmental differences in source monitoring provide one way of clarifying such inquiries as well as determining the basic mechanisms relevant to memory for source and, in turn, may illuminate important aspects of children's memory development in general (Lindsay et al., 1991).

References

Bem, S. (1983). Gender schema theory and its implications for child development: Raising gender-aschematic children in a gender-schematic society. Signs: Journal of Women in Culture and Society, 8(4), 598-616.

Bright-Paul, A., Jarrold, C., and Wright, D. B. (2008). Theory-of-mind development influences suggestibility and source monitoring. *Developmental Psychology*. 44, 1055–1068. doi: 10.1037/00121649.44.4.1055

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. Psychological Review, 106(4), 676-713.

Case, R. (1985). *Intellectual development: Birth to adulthood*. New York: Academic Press.

Case, R., Kurland, M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, 33, 386–404.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: Lawerence Erlbaum.

Day, K., Howie, P., & Markham, R. (1998). The role of similarity in developmental differences in source monitoring. *British Journal of Developmental Psychology*, 16, 219–232.

De Luca, C. R., and Leventer, R. J. (2008). *Developmental Trajectories of Executive Functions Across the Lifespan.* Philadelphia, PA: Taylor & Francis.

Drewe, E. (1975). Go-no-go learning after frontal lobe lesion in humans. *Cortex, 11,* 8-16.

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

Ferguson, S., Hashtroudi, S., & Johnson, M. (1992). Age differences in using source-relevant cues. *Psychology and Aging, 7*(3), 443-452.

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

Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined actions: A developmental comparison. *Child Development*, 56,1145–11 55.

Foley, M. A., Wilder, A., McCall, R., & Van Vorst, R. (1993). The consequences for recall of children's ability to generate interactive imagery in the absence of external supports. *Journal of Experimental Child Psychology*, 56,173–200.

Gerrie, M. P., and 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

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

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

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

Lorsbach, T.C., & Reimer, J.F. (2005). Feature binding in children and young adults. *The*

Journal of Genetic Psychology, 166, 313-327.

Luria, A. R., Pribram, K. H., Homskaya, E. D. (1964). An experimental analysis of the behavioral disturbance produced by a left frontal arachnoidal endothelioma (meningioma). *Neuropsychologia*, *19*, 65-71. doi: 10.1016/0028-3932(64)90034-X

Metcalfe, J., Mencl, W. E., & Cottrell, G. W. (1994). Cognitive binding. In D. L. Schacter & E. Tulving (Eds.), *Memory systems* (pp. 369–394). Cambridge, MA: MIT Press.

Powell, M. B., Roberts, K. P., Ceci, S. J., & Hembrooke, H. (1999). The effects of repeated experience on children's suggestibility. Developmental Psychology, 35, 1462–1477.

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.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515-526.

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, 403-43.

Roberts, K. P., & Blades, M. (1995). Children's discriminations of memories for actual and pretend actions in a hiding task. *British Journal of Developmental Psychology*, 13, 321–334.

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

Roberts, K. P., & Blades, M. (1999). Children's memory and source monitoring for real-life and television events. *Journal of Applied Developmental Psychology*, 20, 575–596.

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

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.

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. doi: 10.1006/jecp.2001.2632

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

Schacter, D. L., Norman, K. A., & Koustaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology*, 49, 289–318.

Thierry, K., & Pipe, M. (2009). The susceptibility of young preschoolers to source similarity effects: Confusing story or video events with reality. *Journal of Experimental Child Psychology*, (102), 392-407.

Thierry, K., & Spence, M. (2002). Source-monitoring training facilitates preschoolers' eyewitness memory performance. *Developmental Psychology*, *38*(3), 428-437.

Welch-Ross, M. K. (1999). Interviewer knowledge and preschoolers' reasoning about knowledge states moderate suggestibility. *Cognitive Development.* 14, 423–442. doi: 10.1016/S08852014(99)00013-1

Welch-Ross, M. K., Diecidue, K., and 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

Table 1

Age, by Age Group

Age Group	N	Mean Age	Standard Deviation	Minimum Age	Maximum Age
3-5	30 (30.3)	5.15	.43	4.54	5.91
6-8	37 (37.4)	7.16	.85	6.01	8.99
18-21	32 (32.3)	19.85	.79	18.30	21.18

Note: Percentage of overall participants in parentheses.

Table 2

Gender, by Age Group

	Age Group		
Gender	3-5	6-8	18-21
Male	11 (36.7)	20 (54.1)	10 (31.3)
Female	19 (63.3)	17 (45.9)	22 (68.8)

Note: Percentage of age group in parentheses.

Table 3

Ethnicity, by Age Group

Age Group	Caucasian/White	Mixed	Minority	"Canadian"	No Response
3-5	7	4	3	5	11
6-8	5	2	2	5	23
18-21	12	2	13	3	2
Total	25	9	18	13	37

Table 4

Mean Overall Source Monitoring Scores for Each Gender, by Age Group

Age Group	Gender	Overall Source Monitoring Score	Standard Error
3-5-Years-Old	Female	.77	.03
	Male	.63	.04
6-8-Years-Old	Female	.81	.03
	Male	.76	.03
18-21-Years-Old	Female	.91	.23
	Male	.87	.04
Overall	Female	.83	.02
	Male	.75	.02

Table 5

Chance Statistics for Recognition of Target Actions, by Age Group

Age Group		Degrees of Freedom	t Value	Significance (One-Tailed)
3-5-Year-Olds	1 Cue	29	7.71	< .001
	5 Cues	29	11.84	< .001
6-8-Year-Olds	1 Cue	36	18.32	< .001
	5 Cues	36	15.25	< .001
18-21-Year-Olds	1 Cue	31	23.28	< .001
	5 Cues	31	16.95	< .001

Table 6

Chance Statistics for Recognition of Distractor Items, by Age Group

Age Group		Degrees of Freedom	t Value	Significance (One-Tailed)
3-5-Year-Olds	1 Cue	29	7.12	< .001
	5 Cues	29	7.86	< .001
6-8-Year-Olds	1 Cue	36	30.16	< .001
	5 Cues	36	19.24	< .001
18-21-Year-Olds	1 Cue	31	56.64	< .001
	5 Cues	31	35.22	< .001

Table 7

Mean Source Monitoring Scores for Each Condition, by Age Group

Age Group	Condition	Source Monitoring Score (Proportion)	Standard Deviation
3-5	1 Cue	.69	.22
	5 Cues	.75	.17
6-8	1 Cue	.74	.16
	5 Cues	.81	.18
18-21	1 Cue	.87	.12
	5 Cues	.92	.12

Note: Main effects of age and cue condition significant at <.001 and .006 respectively.

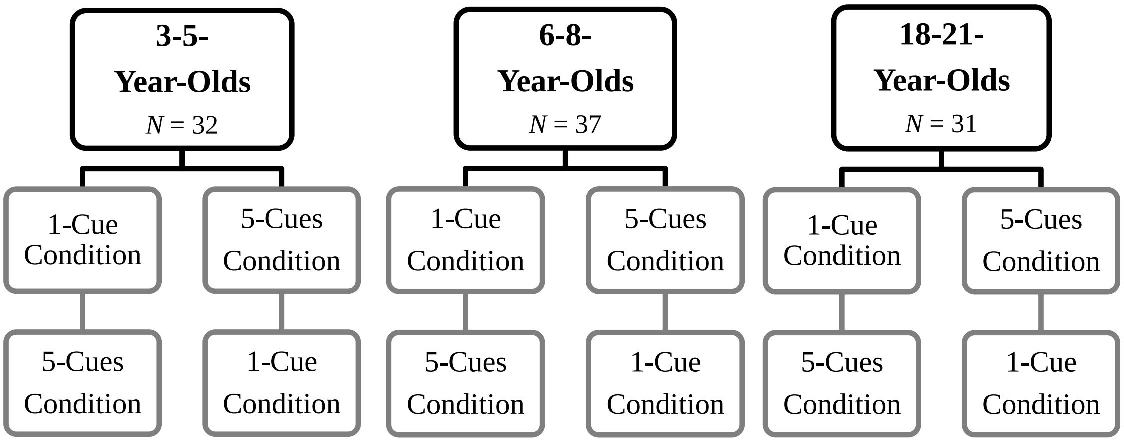

Figure 1. Study design including age group and condition.

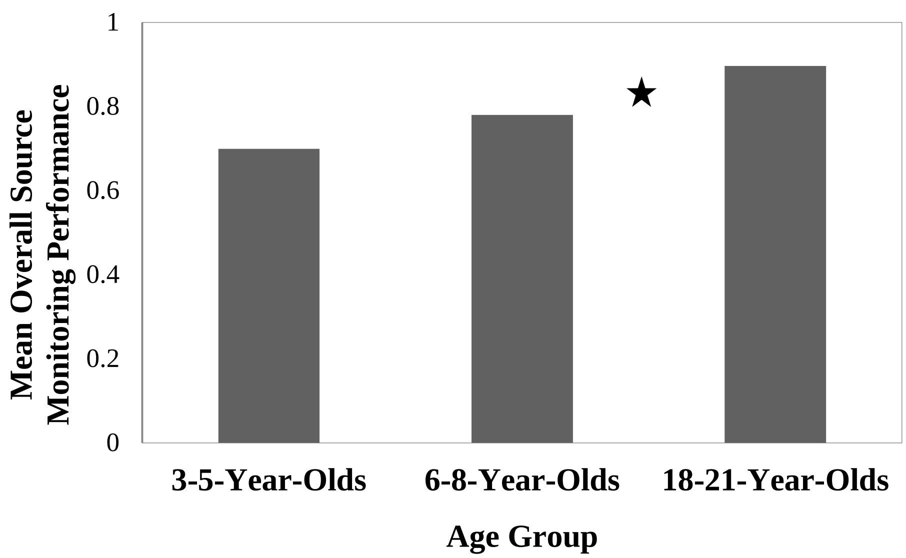

Figure 2. Graphical results of hypothesis 1: An overall/general developmental progression in participants' ability to make accurate source judgments.

Appendix A

Note: 1 Cue Condition; 1 salient difference between sources (t-shirt colour).

Note: 5 Cues Condition; 5 differences between sources (presence/non-presence of hat, hair colour, presence/non-presence of necklace, t-shirt colour, pant colour).

Appendix B

Action Lists

	Action List 1	**Action List 2**	**Action List 3**	**Action List 4**
Target Actions	Clap hands 3 times	Recite the days of the week	"I love dogs"	Wave flag from side-to-side (6 times)
	"I'm going to eat an apple"	Toss and catch a ball 3 times	Jump 3 times	Close eyes (5 seconds)
	Hands on hips (5 seconds)	Rub belly in circular motion with right hand (5 seconds)	"Rainbows are (recite the colours of the rainbow: red, orange, yellow, green, blue, indigo, violet)"	March on the spot (5 seconds)
	"What time is it?"	"Math is fun!"	Take a sip of water	Answer phone ("Hello?")
	Look up and down 3 times (pausing at center)	Shake Tambourine (5 seconds)	Tap head 3 times with right hand	Turn 360 (clock-wise) on the spot 3 times
	Hula-hoop	"My favourite colour is green"	Rip a piece of paper in half	"No"
Distractor Items	Blow whistle	Jump 3 times with a skipping rope	Take a bite of a banana	Sit down on the ground and stand back up
	Cover eyes with both hands	Two thumbs up	Cover mouth with right hand and yawn	"I don't like snakes"
	"Let's go to the beach!"	"Yes"	Wave at camera	Recite the months of the year
	Blow up a balloon	Sneeze	Cover ears with both hands	Put blanket over shoulders ("burr")

Note: Distractor items are not actually performed by actors—they are non-present items that participants are asked recognition questions is regards to, for each action list.

Appendix C

Video Forms

Counterbalancing of Conditions, Source Presentation and Corresponding Action Lists

Video Form	Condition Presentation	Source Order	Action List Order
Video 1	Condition 1	Abby 1st Paige 2nd	Action List 1 Action List 2
	Condition 2	Alexia 1st Candice 2nd	Action List 3 Action List 4
Video 2	Condition 2	Alexia 1st Candice 2nd	Action List 4 Action List 3
	Condition 1	Abby 1st Paige 2nd	Action List 2 Action List 1
Video 3	Condition 1	Paige 1st Abby 2nd	Action List 2 Action List 1
	Condition 2	Candice 1st Alexia 2nd	Action List 4 Action List 3
Video 4	Condition 2	Candice 1st Alexia 2nd	Action List 3 Action List 4
	Condition 1	Paige 1st Abby 2nd	Action List 1 Action List 2

Note: Condition 1 = 1 Cue Condition, Sources: Abby & Paige

Condition 2 = 5 Cues Condition, Sources: Candice & Alexia

Appendix D

Adult Distractor Task

The History of the Chocolate Chip Cookie

Cookie lovers around the world have been enjoying the famous chocolate chip cookie for decades. There is no sign of scarcity for the much-desired cookie in any kitchen. People of all ages swarm for the delicious treat; it's safe to say that no pantry, cupboard, or stove is safe. But where did the chocolate chip cookie come from? Who created this delicious combination of flavours that has stimulated the senses of people for generations, and proven to be a tradition in the hearts of many families?

It was Ruth Graves Wakefield, a restaurant owner and cook in Whitman, Massachusetts. Rumor has it that in 1930, Wakefield was mixing a batch of cookies when she discovered that she was out of baker's chocolate. She substituted broken pieces of semi-sweet chocolate, expecting it to melt and absorb into the dough to create chocolate cookies. When the timer went off, the cookies were cooled, and the children at her restaurant were lining up for a treat, Wakefield discovered she had stumbled upon a masterpiece. She called it the 'Toll House Cookie' (named after her acclaimed restaurant) until the typical American name took over.

Today there are too many variations of the chocolate chip cookie to count. New flavours such as M & M's, macadamia nut, and chocolate peanut butter have sprouted from the original chocolate chip cookie's historic soil. Although these competitors put up a good battle, even today, there truly is none like the original chocolate chip cookie.

So, thank you Mrs. Wakefield! You are the one who truly makes good cookies!

Appendix E

Source Monitoring Interview Response Sheets

Response Sheet Video 1

	Action	Answer	If yes, who?
SEGMENT 1			
	Did someone look up and down 3 times?	Yes No	
	Did someone rub their belly in a circular motion?	Yes No	
	Did someone put their hands on their hips?	Yes No	
	Did someone say "Let's go to the beach!"?	Yes No	
	Did someone say "What time is it?"?	Yes No	
	Did someone jump 3 times with skipping rope?	Yes No	
	Did someone blow up a balloon?	Yes No	
	Did someone say "Math is fun!"?	Yes No	
	Did someone give two thumbs up?	Yes No	
	Did someone say the days of the week?	Yes No	
	Did someone say "I'm going to eat an apple"?	Yes No	
	Did someone use a hula-hoop?	Yes No	
	Did someone blow a whistle?	Yes No	
	Did someone shake a tambourine?	Yes No	
	Did someone say "Yes"?	Yes No	
	Did someone sneeze?	Yes No	
	Did someone toss and catch a ball 3 times?	Yes No	
	Did someone say "My favourite colour is green"?	Yes No	
	Did someone clap their hands 3 times?	Yes No	
	Did someone cover their eyes with both hands?	Yes No	

SEGMENT 2

	Did someone wave a flag from side-to-side?	Yes No	
	Did someone turn around on the spot 3 times?	Yes No	
	Did someone recite the months of the year?	Yes No	
	Did someone say "No"?	Yes No	
	Did someone put a blanket over their shoulders and say "burr"?	Yes No	
	Did someone say "I don't like snakes"?	Yes No	
	Did someone cover their mouth with their hand and yawn?	Yes No	
	Did someone take a sip of water?	Yes No	
	Did someone march on the spot?	Yes No	
	Did someone rip a piece of paper in half?	Yes No	
	Did someone answer the phone and say "Hello"?	Yes No	
	Did someone jump up and down 3 times?	Yes No	
	Did someone sit down on ground and stand back up?	Yes No	
	Did someone cover their ears with both hands?	Yes No	
	Did someone close their eyes for 5 seconds?	Yes No	
	Did someone tap their head 3 times with their right hand?	Yes No	
	Did someone wave their hand?	Yes No	
	Did someone take a bite of a banana?	Yes No	
	Did someone say the colours of the rainbow?	Yes No	
	Did someone say "I love dogs"?	Yes No	

Which was harder…

Remembering what Abby & Paige did or what Candice & Alexia did?	Abby & Paige	Candice & Alexia

Why?

Response Sheet Video 2

SEGMENT 1	Action	Answer	If yes, who?
	Did someone answer the phone and say "Hello?"	Yes No	
	Did someone take a sip of water?	Yes No	
	Did someone say "I don't like snakes"?	Yes No	
	Did someone wave their hand?	Yes No	
	Did someone say the months of the year?	Yes No	
	Did someone tap their head 3 times with their right hand?	Yes No	
	Did someone march on the spot?	Yes No	
	Did someone say the colours of the rainbow?	Yes No	
	Did someone take a bite of a banana?	Yes No	
	Did someone cover their ears with both hands?	Yes No	
	Did someone turn around on the spot 3 times?	Yes No	
	Did someone wave a flag from side-to-side?	Yes No	
	Did someone say "No"?	Yes No	
	Did someone put a blanket over their shoulders and say "burr"?	Yes No	
	Did someone close their eyes for 5 seconds?	Yes No	
	Did someone cover their mouth with their hand and yawn?	Yes No	
	Did someone say "I love dogs"?	Yes No	
	Did someone sit down on ground and stand back up?	Yes No	
	Did someone rip a piece of paper in half?	Yes No	
	Did someone jump up and down 3 times?	Yes No	

SEGMENT 2

	Did someone clap their hands 3 times?	Yes No	
	Did someone sneeze?	Yes No	
	Did someone look up and down 3 times?	Yes No	
	Did someone give two thumbs up?	Yes No	
	Did someone toss and catch a ball 3 times?	Yes No	
	Did someone say "Math is fun!"?	Yes No	
	Did someone say the days of the week?	Yes No	
	Did someone blow up a balloon?	Yes No	
	Did someone say "What time is it?"?	Yes No	
	Did someone say "My favourite colour is green"?	Yes No	
	Did someone jump 3 times with skipping rope?	Yes No	
	Did someone say "Let’s go to the beach!"?	Yes No	
	Did someone cover their eyes with both hands?	Yes No	
	Did someone say "Yes"?	Yes No	
	Did someone use a hula-hoop?	Yes No	
	Did someone shake a tambourine?	Yes No	
	Did someone rub their belly in a circular motion?	Yes No	
	Did someone say "I'm going to eat an apple"?	Yes No	
	Did someone put their hands on their hips?	Yes No	
	Did someone blow a whistle?	Yes No	

Which was harder…

Remembering what Abby & Paige did or what Candice & Alexia did?	Abby & Paige	Candice & Alexia

Why?

Response Sheet Video 3

SEGMENT 1	**Action**	**Answer**	**If yes, who?**
	Did someone clap their hands 3 times?	Yes No	
	Did someone say "Let's go to the beach!"?	Yes No	
	Did someone look up and down 3 times?	Yes No	
	Did someone say "I'm going to eat an apple"?	Yes No	
	Did someone use a hula-hoop?	Yes No	
	Did someone say the days of the week?	Yes No	
	Did someone shake a tambourine?	Yes No	
	Did someone say "My favourite colour is green"?	Yes No	
	Did someone jump 3 times with skipping rope?	Yes No	
	Did someone say "Math is fun!"?	Yes No	
	Did someone cover their eyes with both hands?	Yes No	
	Did someone put their hands on their hips?	Yes No	
	Did someone say "Yes"?	Yes No	
	Did someone toss and catch a ball 3 times?	Yes No	
	Did someone blow up a balloon?	Yes No	
	Did someone say "What time is it?"?	Yes No	
	Did someone rub their belly in a circular motion?	Yes No	
	Did someone give two thumbs up?	Yes No	
	Did someone blow a whistle?	Yes No	
	Did someone sneeze?	Yes No	

SEGMENT 2

	Did someone say the colours of the rainbow?	Yes No	
	Did someone sit down on ground and stand back up?	Yes No	
	Did someone tap their head 3 times with their right hand?	Yes No	
	Did someone take a sip of water?	Yes No	
	Did someone put a blanket over their shoulders and say "burr"?	Yes No	
	Did someone rip a piece of paper in half?	Yes No	
	Did someone wave their hand?	Yes No	
	Did someone take a bite of a banana?	Yes No	
	Did someone answer the phone and say "Hello"?	Yes No	
	Did someone cover their mouth with their hand and yawn?	Yes No	
	Did someone say "No"?	Yes No	
	Did someone turn around on the spot 3 times?	Yes No	
	Did someone say "I don't like snakes"?	Yes No	
	Did someone wave a flag from side-to-side?	Yes No	
	Did someone cover their ears with both hands?	Yes No	
	Did someone say "I love dogs"?	Yes No	
	Did someone march on the spot?	Yes No	
	Did someone close their eyes for 5 seconds?	Yes No	
	Did someone jump up and down 3 times?	Yes No	
	Did someone say the months of the year?	Yes No	

Which was harder…

Remembering what Abby & Paige did or what Candice & Alexia did?	Abby & Paige	Candice & Alexia

Why?

Response Sheet Video 4

	Action	Answer	If yes, who?
SEGMENT 1			
	Did someone cover their mouth with their hand and yawn?	Yes No	
	Did someone wave a flag from side-to-side?	Yes No	
	Did someone say the colours of the rainbow?	Yes No	
	Did someone turn around on the spot 3 times?	Yes No	
	Did someone say "I love dogs"?	Yes No	
	Did someone rip a piece of paper in half?	Yes No	
	Did someone say the months of the year?	Yes No	
	Did someone answer the phone and say "Hello"?	Yes No	
	Did someone sit down on ground and stand back up?	Yes No	
	Did someone jump up and down 3 times?	Yes No	
	Did someone close their eyes for 5 seconds?	Yes No	
	Did someone march on the spot?	Yes No	
	Did someone take a bite of a banana?	Yes No	
	Did someone say "No"?	Yes No	
	Did someone cover their ears with both hands?	Yes No	
	Did someone put a blanket over their shoulders and say "burr"?	Yes No	
	Did someone wave their hand?	Yes No	
	Did someone tap their head 3 times with their right hand?	Yes No	
	Did someone say "I don't like snakes"?	Yes No	
	Did someone take a sip of water?	Yes No	

SEGMENT 2

	Did someone put their hands on their hips?	Yes No	
	Did someone give two thumbs up?	Yes No	
	Did someone clap their hands 3 times?	Yes No	
	Did someone sneeze?	Yes No	
	Did someone say "What time is it?"?	Yes No	
	Did someone use a hula-hoop?	Yes No	
	Did someone say "Yes"?	Yes No	
	Did someone say "Let's go to the beach!"?	Yes No	
	Did someone toss and catch a ball 3 times?	Yes No	
	Did someone say the days of the week?	Yes No	
	Did someone cover their eyes with both hands?	Yes No	
	Did someone look up and down 3 times?	Yes No	
	Did someone rub their belly in a circular motion?	Yes No	
	Did someone say "My favourite colour is green"?	Yes No	
	Did someone jump 3 times with skipping rope?	Yes No	
	Did someone say "Math is fun!"?	Yes No	
	Did someone blow a whistle?	Yes No	
	Did someone say "I'm going to eat an apple"?	Yes No	
	Did someone shake a tambourine?	Yes No	
	Did someone blow up a balloon?	Yes No	

Which was harder…

Remembering what Abby & Paige did or what Candice & Alexia did?	Abby & Paige	Candice & Alexia

Why?

Appendix F

Verbal Assent Script

“Hi! My name is __________, and I work at a university! Do you know what a university is? It’s a really big school. I’ve come here today to do a fun activity with you. I’ve called it the ‘Laurier activity’ because some people at a place called Laurier University helped me to get all the things ready for what we are going to do today. I think we’re going to have a lot of fun. And if you ever don’t want to carry on, you can just tell me and I’ll take you back to your classroom.”

“Does this sound like something you would like to do?”

If no: “That’s okay. I’ll take you back to your classroom.”

If yes: “Great. Let’s go!”

Appendix G

Actor Headshots

Appendix H

Coding Sheets

Coding Form- Video 1

Segment 1 (one-cue condition):

	Total		
Number of target recognition questions *(i.e., yes/no*) correct out of 12		*Total when correct source was Abby*:	*Total when correct source was Paige:*
Number of target sources *(i.e., Abby/Paige)* correct out of number of target recognition questions correct *(above)*		*Total- Abby:*	*Total- Paige:*
Number of distractors correct out of 8			
Distractor Source- Abby			
Distractor Source- Paige			

Segment 2 (five-cues condition):

	Total		
Number of target recognition questions *(i.e.., yes/no)* correct out of 12		*Total when correct source was Candice*:	*Total when correct source was Alexia*:
Number of target sources *(i.e., Candice/Alexia)* correct out of number of target recognition questions correct *(above)*		*Total- Candice:*	*Total- Alexia:*
Number of distractors correct out of 8			
Distractor Source- Candice			
Distractor Source- Alexia			

Coding Form- Video 2

Segment 1 (five-cues condition):

	Total		
Number of target recognition questions *(i.e.. yes/no)* correct out of 12		*Total when correct source was Candice*:	*Total when correct source was Alexia*:
Number of target sources *(i.e., Candice/Alexia)* correct out of number of target recognition questions correct *(above)*		*Total- Candice:*	*Total- Alexia:*
Number of distractors correct out of 8			
Distractor Source- Candice			
Distractor Source- Alexia			

Segment 2 (one-cue condition):

	Total		
Number of target recognition questions *(i.e., yes/no*) correct out of 12		*Total when correct source was Abby*:	*Total when correct source was Paige:*
Number of target sources *(i.e., Abby/Paige)* correct out of number of target recognition questions correct *(above)*		*Total- Abby:*	*Total- Paige:*
Number of distractors correct out of 8			
Distractor Source- Abby			
Distractor Source- Paige			

Coding Form- Video 3

Segment 1 (one-cue condition):

	Total		
Number of target recognition questions *(i.e., yes/no*) correct out of 12		*Total when correct source was Abby*:	*Total when correct source was Paige:*
Number of target sources *(i.e., Abby/Paige)* correct out of number of target recognition questions correct *(above)*		*Total- Abby:*	*Total- Paige:*
Number of distractors correct out of 8			
Distractor Source- Abby			
Distractor Source- Paige			

Segment 2 (five-cues condition):

	Total		
Number of target recognition questions *(i.e.. yes/no)* correct out of 12		*Total when correct source was Candice*:	*Total when correct source was Alexia*:
Number of target sources *(i.e., Candice/Alexia)* correct out of number of target recognition questions correct *(above)*		*Total- Candice:*	*Total- Alexia:*
Number of distractors correct out of 8			
Distractor Source- Candice			
Distractor Source- Alexia			

Coding Form- Video 4

Segment 1 (five-cues condition):

	Total		
Number of target recognition questions *(i.e.. yes/no)* correct out of 12		*Total when correct source was Candice*:	*Total when correct source was Alexia*:
Number of target sources *(i.e., Candice/Alexia)* correct out of number of target recognition questions correct *(above)*		*Total- Candice:*	*Total- Alexia:*
Number of distractors correct out of 8			
Distractor Source- Candice			
Distractor Source- Alexia			

Segment 2 (one-cue condition):

	Total		
Number of target recognition questions (*i.e., yes/no*) correct out of 12		*Total when correct source was Abby*:	*Total when correct source was Paige:*
Number of target sources *(i.e., Abby/Paige)* correct out of number of target recognition questions correct *(above)*		*Total- Abby:*	*Total- Paige:*
Number of distractors correct out of 8			
Distractor Source- Abby			
Distractor Source- Paige			

Appendix I

Common Reasons for Subjective Difficulty

	One-Cue	Five-Cues
3-5-Year-Olds	- "Had to think hard" - "Couldn't remember" - "Because of [specific actions]" - "I don't know" - "They did harder things"	– "I don't know" – "Couldn't remember – "Harder to remember"
6-8-Year-Olds	– "I don't know" – "They did harder things" – Reference to order (e.g., "they were first/last")	– "The others were easier" – "I don't know" – References to differences in clothing (hat, necklace) – Reference to order
18-21-Year-Olds	– Reference to order – Reference to similarity/dissimilarity of clothing and features – Found dissimilar easier to differentiate/more memorable	– Their clothing made them easier – Reference to order – Their actions were more memorable