2007

Interviewing Children about a Repeated Event: Does Prior Practice in Describing a Specific Instance of an Unrelated Repeated Event Improve the Amount and Quality of Elicited Information?

Sonja P. Brubacher
Wilfrid Laurier University

Interviewing Children about a Repeated Event: Does Prior Practice in Describing a Specific Instance of an Unrelated Repeated Event Improve the Amount and Quality of Elicited Information?

by

Sonja P. Brubacher

Honours Bachelor of Arts, Wilfrid Laurier University, 2004

THESIS

Submitted to the Department of Psychology

In partial fulfilment of the requirements for

Master of Arts, Social & Developmental Psychology

Wilfrid Laurier University

2007

Your file Votre référence
ISBN: 978-0-494-32270-3
Our file Notre référence
ISBN: 978-0-494-32270-3

Abstract

Five- to 6- and 7- to 8-year olds ($N = 118$) participated, in groups, in 4 sessions of the Laurier Activities over a 2-week period, The Laurier Activities centred around tasks such as listening to a story, completing a puzzle, mild physical exercise, relaxation, getting refreshed, and getting a surprise. Across sessions of the activities, instantiations of each task (e.g., the *content of the story*) were presented at different frequencies. Instantiations were fixed (the same every time), variable (changed every time), or "Hi/Lo" (the Hi frequency instantiation presented at 3 sessions, the Lo frequency instantiation presented at 1 session). Children were interviewed individually based on the NICHD protocol (Orbach et al., 2000) after a week delay. This protocol consists of rapport-building, ground rules, a practice phase in which the child is encouraged to report episodic details in response to open-ended questions, and a substantive phase. The current study focused on the Practice Phase of this interview and manipulated the type of practice in 3 between-subjects conditions. In the experimental conditions, children practiced describing *specific instances* of a repeated autobiographical event (**incident-specific practice**), or what generally happens during the repeated event (**general practice**). Children in the **control** condition described a single-experience, novel, event. Incident-specific practice benefited children most when recalling the Laurier Activities by encouraging them to disclose multiple incidents earlier, to recall more details, and to mention more differences across occurrences. Age by condition interactions suggested, however, that incident-specific practice led to a larger increase in performance over other conditions for younger children, while older children tended not to differ significantly across conditions on many variables of interest. Differences in accuracy across conditions were not significant for

either age group. Findings are discussed in terms of practical implications for field interviewers.

Acknowledgments

I would like to thank my advisor, Dr. Kim Roberts, for her enduring support throughout this project, as well as her patience, insights, and encouragement over the years that I have had the privilege to work with her. She has been an inspiring mentor, and I look forward to continuing our work together into the future.

Additionally, I would like to thank the members of my committee, Dr. Alexandra Gottardo, and Dr. Michael Pratt, for their willingness to advise on this project, their helpful insights, and for their valuable time.

The research presented in this thesis was carried out in local schools and daycares, as well as in the Child Memory and Learning Laboratory at Wilfrid Laurier University. I would like to thank members of this lab, as well as honorary members, for their help in collecting data, interviewing children, transcribing, and coding. Each one of them played a role in making this project possible. Thank you to Valerie Vorstenbosch, Lisa Gravel, Leanne Best, Ashley Graham, Alana Greco, Susan Battista, Tara Dumas, Nicole Keir, Taryn Moss, Una Glisic, Heather Gough, Courtney Patterson, Melanie Crutchley, and Brian Hoessler. Furthermore, I am indebted to the parents, children, teachers, principals, and daycare workers who made data collection possible. Additional thanks to the Natural Sciences and Engineering Research Council of Canada (NSERC), for a Postgraduate Scholarship – Masters (PGS – M), which provided me with financial support while I worked towards earning my Masters degree.

Finally, I am grateful for the support of my family, my partner, and my friends. They always believed I would succeed, and provided endless encouragement.

Table of Contents

List of Tables

Many children who make allegations of sexual abuse have experienced the abuse on repeated occasions. For example, in a representative sample of 98 children drawn from a pool of over 1000 interviews with child sexual abuse victims in Israel, 42% of the 98 cases involved three or more instances of abuse (Lamb, Sternberg, Esplin, Hershkowitz, Orbach & Hovav, 1997). Children's reports of repeated events are qualitatively different from their reports of novel, single-experience events (see Roberts & Powell, 2001, for a review). While children who have experienced an event multiple times are highly accurate about details that are constant across occasions, they are less accurate about details that are variable, and are more likely to confuse these variable details across occurrences (Powell, Roberts, Ceci, & Hembrooke, 1999). Many techniques are currently being researched to help prevent children making these between-event confusions (e.g., source-monitoring training studies; Poole & Lindsay, 2001; Thierry & Spence, 2002) because when children are asked to testify about a repeated event, they are usually required to describe one or two instances with a reasonable amount of precision, such as providing time, place, and the actions that occurred, and thus must be able to discriminate among different occasions (*S v. R,* 1989, as cited in Powell & Thomson, 2003; Roberts, 2002). Without this information available, it is difficult for the accused to refute the allegations (e.g., without date and time information, there is no opportunity to provide an alibi) (Guadagno & Powell, 2006).

Additionally, interviewing protocols designed to elicit the most accurate information from children are becoming increasingly more grounded in theoretical frameworks. The protocol developed by researchers at the National Institute of Child Health and Human Development (NICHD; Orbach, Hershkowitz, Lamb, Sternberg,

Esplin, & Horowitz, 2000) enforces the use of invitations and open-ended, non-suggestive questions, aimed at eliciting the most complete accounts from children while preserving accuracy. While several studies have demonstrated that this interview technique increases the quantity of information provided, both in samples of children who have experienced abuse one time and those who have had multiple experiences, data from these groups have not been directly compared (Orbach et al., 2000; Sternberg, Lamb, Orbach, Esplin & Mitchell, 2001). It is currently not known if the structure of the protocol is equally appropriate for children with single and multiple experiences of abuse. Furthermore, the majority of research upon which these interview guidelines were developed has been focused on children's memories for, and subsequent reporting of, single events.

Following a short rapport-building phase in which the child becomes comfortable with the interviewer, and prior to the substantive (target) phase, the NICHD protocol recommends a "presubstantive" practice phase in which the child and interviewer engage in a discussion of a neutral past event. This phase allows the child to practice responding to open-ended prompts, and to understand the type of communication that will be expected in the substantive phase (Orbach et al., 2000). It is suggested that the neutral event be a holiday (e.g., Christmas); however, if a holiday has not occurred recently, another neutral event will be queried. The type of event discussed at this phase could potentially have a number of different characteristics with respect to its frequency. A holiday is a repeated event that happens on a yearly basis. It would also be possible, however, to ask a child to describe a more frequent repeated event, such as weekly swimming lessons, or about a novel event, such as a recent first trip to a museum.

Although one goal of this phase is to encourage the child to provide incident-specific "episodic" details (e.g., what happened *last* Halloween), if a repeated event serves as the practice topic, the child's narrative may continue to include scripted details. Conversely, the child may provide details specific to a particular instance (e.g., the *last time*) if the interviewer is sensitive to script-like language in the child's dialogue (e.g., "we *get* candy" versus "we *got* candy"), and thus makes explicit requests to the child to describe a specific instance. The question then arises as to whether it might be effective to practice describing certain types of events (i.e., repeated) and whether the quality of the child's substantive narrative can be enhanced by explicitly directed practice in describing a *specific instance of a repeated event*. In addition, new directions in research on the current protocol interview are focusing on other aspects of the interview than just the nature and type of questions being asked in the substantive phase. For example, new studies are focusing on the effects of interviewer paraphrasing of child utterances (Evans & Roberts, 2007), which could have both social and cognitive benefits, as well as social implications of the rapport building phase (McCauley, Langrock, & Redmond, 2007, April).

Furthermore, while research exists on both the rapport building phase, and the substantive portions of the NICHD protocol, it has never been examined whether the episodically-oriented Practice Phase is actually effective in demonstrating to children how they are to answer questions in the substantive phase.

Goals of Current Study

The aim of the current research was to compare the quality of children's narratives for a laboratory repeated event when the children had engaged in three types of

practice conditions. One third of the children in the study practiced recalling two specific instances of a repeated event from their daily lives (**incident-specific practice**), while another third of the children practiced describing what "usually happens" when they engaged in a similar autobiographical repeated event (**general practice**). The final group of children served as a **control** group, and practiced describing a recent novel (i.e., single-experience) event.

When children are asked to recall a repeated event, without any particular instructions, their narratives typically consist of a mix of scripted information (e.g., what usually happens), as well as some incident-specific details (Farrar & Boyer-Pennington, 1999). Such narratives can be problematic if a child is testifying about a serious event, such as abuse. While a few jurisdictions allow children who have been exposed to repeated abuse to testify in a general manner about what usually happens, the norm is that children are required to give specific details about at least one or two events (Powell & Thomson, 2003). Before moving to the specific hypotheses of the current study, our understanding of children's memories of repeated events will be reviewed by discussing three major theoretical frameworks. First, script theories (e.g., Farrar & Goodman, 1992; Nelson, 1986) will be reviewed, followed by Fuzzy-Trace Theory (Brainerd & Reyna, 1990), and lastly the source-monitoring framework will be discussed (Johnson, Hashtroudi, & Lindsay, 1993). Finally, contributions from these three frameworks to the understanding of children's memories for single experiences will be very briefly reviewed, as the current study does not focus on memory for novel events. An appropriate control condition, to compare with practice in describing repeated events, was to have children describe novel events.

Memory for repeated events

Script Theories. The use of scripts to aid memory retrieval is a phenomenon occurring across the lifespan (Nelson, 1986). Scripts can be conceptualized as cognitive structures which help people organize and make sense of common experiences (Abelson, 1981) such as what happens at a restaurant or a birthday party. With increasing experience, these structures become more flexible, such that one's script for going to a restaurant could allow inclusion of both ordering food at a counter or ordering from a waiter at a table. "Ordering food" is represented in the generic script, with a "slot" into which any number of script-consistent activities (e.g., ordering at a counter, from a waiter) can fit, and are stored as a list (Nelson, 1986). Scripts are also organized temporally (Mandler, 1979), such that certain elements (e.g., ordering food) will precede, and be related to, other elements of the script (e.g., eating the food).

Children as young as 3-years old will describe highly familiar events, such as what happens at a restaurant, in an organized, schema-dependent manner, although their narratives include relatively few details (Hudson, Fivush, & Kuebli, 1992). In fact, children's scripts of these highly familiar events can be so well developed that they may report having experienced central script details even when they have not occurred. Erskine, Markham and Howie (2002) showed 5- to 6- and 9- to 10-year old children slideshows of a trip to McDonalds. One slide sequence omitted the three most central details that children had mentioned during a pilot study as being part of a trip to McDonalds (e.g., paying for the food). The 5- to 6-year old children were much more likely than the older children to inaccurately report having seen the central details.

Other researchers have demonstrated similar findings for the prevalence of script-based memories in young children when events are highly familiar. For example, when asked about having dinner at home (Hudson, 1983; as cited in Fivush, Hudson, & Nelson, 1984), or about kindergarten (Fivush, 1982; as cited in Fivush et al., 1984), 3- to 5-year olds found it much easier to describe what usually happens, than to describe what happened the last time they engaged in the activity.

What can be expected from children's memories of less familiar events? In a study by Fivush and colleagues (1984), children were asked what usually happens when they go to a museum. Two weeks later, the children went to a new museum, engaged in some novel activities (such as digging for "artefacts"), and several interviews were conducted at later time periods. Half of the children were questioned about what "happens when you go to a museum" while the other half of the children were explicitly questioned about what "happened" at the Jewish museum, which contained quite novel experiences. While the narratives of these 5-year olds still tended to contain some script-like language (e.g., present tense: "*you do* X"), children who were asked specifically about the Jewish museum used past tense language and reported specific details about the trip. A few of these children were interviewed one year after the event, and although they remembered fewer details, their narratives continued to be event specific, and well structured.

Related research on children's script memory has examined how children of different ages process events that differ slightly from an original event, but remain consistent with the general event schema. For example, Farrar and Goodman (1992) engaged 4- and 7-year old children in a standard event either once or three times, and

then presented them with an event in which some of the details varied (deviation event). The younger children who participated three times were more likely to rely on the general script that they had developed for the standard event, than were the older children, and thus they incorporated details from the deviation event into their memories of the standard event. The 7-year olds were better able to form distinct memories for the standard and deviation visits. Interestingly, children of both age groups had strong memories for the standard event regardless of whether they had experienced it one time (with no deviation event), or three times (with one deviation event following the final standard event). Children were most confused when they had experienced one standard and one deviation event.

The schema-confirmation-deployment model (Farrar & Goodman, 1992) provides some insight as to why the younger children who participated three times, in Farrar and Goodman's study, confused details across occurrences, whereas the older children were less confused, even though children of both ages had built a script for the standard event. The basic assumption of this model is that a schema for an event first must be built. Once a schema exists, similar schema-consistent events can be compared and the schema undergoes further *confirmation*. Younger children take longer at the confirmation phase than do older children. During this phase, any schema-consistent details tend to be absorbed into the script. Only once the deployment phase is reached do children begin to recognize which details *always* belong to the script and which details only belong *sometimes* (Farrar & Goodman, 1992).

According to the schema-confirmation-deployment model (Farrar & Goodman, 1992), young children take longer to build scripts, and therefore, when they are

confronted with a fairly novel event, much of their cognitive resources are devoted to script building and they may either not notice deviations or may simply incorporate them into their generic script. Only at the deployment phase, when scripts are well-developed, do children begin to compare deviations with their script memory. Older children progress through the schema-confirmation phase much more quickly and are thus able to recognize deviations sooner. Additionally, the more similar instances of a repeated event are to one another, the less likely it is that separate episodic memories will be retrieved, by people of any age. That is, major event deviations result in the formation of a separate memory, while deviations that are minor and script consistent are more likely to be "absorbed" into the script (Farrar & Goodman, 1992). Such a finding suggests that when considering similar instances of repeated events it may be very difficult to remember the source of various details that are present during some occurrences and not others. It should be noted, however, that the schema-confirmation-deployment model has been used to suggest that, under some circumstances, if children eventually reach the deployment stage they *may* be able to retrospectively identify deviations and modify their scripts (Farrar & Goodman, 1992).

In general, script theory is useful for predicting what will be remembered about a repeated event (i.e., major event details that occur every time and are central to the script), however, it is less precise with respect to overcoming source confusions. Because the theory suggests that minor deviations in the script may be completely overlooked (i.e., not encoded as a deviation) or absorbed into the script (a failure to keep details separate in storage), there is no means by which to elicit an accurate retrieval of the surface details of any one instance of the event.

Fuzzy-trace theory. Fuzzy-trace theory (Brainerd & Reyna, 1990; Brainerd & Reyna, 1993; Brainerd & Reyna, 1998; Brainerd & Reyna, 2004) is a dual-process model of cognition in that general event representations ("gist traces") are encoded and stored independently and in parallel with surface details ("verbatim traces"). This theory would suggest that any details that are consistent with the overall theme of the event ('gist consistent') may be assigned to any instance of the repeated event, if the exact surface representations of each instance ('verbatim traces') have decayed (Brainerd & Reyna, 2004).

Fuzzy-trace theory (Brainerd & Reyna, 1990; Brainerd & Reyna, 1993; Brainerd & Reyna, 1998; Brainerd & Reyna, 2004) also provides a theoretical account of how children make source confusions in their recall of repeated events. According to this theory, source is simply another detail that is encoded, and thus can be remembered, or conversely, forgotten. For any given experience, two traces are encoded separately and in parallel. The verbatim trace stores all of the surface details of the experience, with source considered to be one of these details. The gist trace contains the overall meaning and structure of the experience. Gist traces for events can be likened to scripts, with each activity of the event being a slot into which any number of gist-consistent verbatim details could be placed. For example, if each session of a repeated experience begins with the child sitting on something, each unique instantiation is a verbatim detail and can be stored in the slot for item-sat-upon. Although there are similarities between script theories and fuzzy-trace theory, there are important differences as well which are highly relevant to the current study. While both theories suggest that general event representations arise from experience ("gists" or "scripts"), fuzzy-trace theory is a *dual*

process model of cognition, allowing for gist and verbatim traces to exist separately (Brainerd & Reyna, 1990). Thus, fuzzy-trace theory permits prediction that a verbatim trace can be later accessed (if it has not completely decayed) because it was encoded separately from the gist. Much research exists supporting this idea of dual traces, including basic cognition studies such as the word-superiority effect (Ankrum & Palmer, 1989; as cited in Brainerd & Reyna, 2004) and the missing letter effect (Moravcski & Healy, 1995), as well as research concerning the development of false memories (Reyna, Holliday & Marche, 2002).

The word superiority effect has been demonstrated through experiments in which a word is presented for a very short duration (less than 100 ms), and then participants are given a yes-no recognition trial, being either a whole word itself, or just one letter. Participants show much better performance in recognizing whole words rather than single letters (Ankrum & Palmer, 1989; as cited in Brainerd & Reyna, 2004). If meaning (gist) were to be built from verbatim traces, rather than these traces being independently processed, one could not recognize a word presented almost below awareness without having identified its constituent letters. The missing letter effect is a similar phenomenon, which has been demonstrated in experiments where participants are asked to circle all instances of a letter in a paragraph (e.g., "t"), and they show a tendency to miss the letter much more often in common words such as "the" than in less common words (Moravcski & Healy, 1995). False memories also support this dual process model of cognition in that details (verbatim information) that are gist-consistent may be falsely recognized as having been present in a particular occurrence, even if they were absent from a to-be-remembered scene (Reyna et al., 2002). For example, if a participant observed a

perpetrator carrying a screwdriver, and was later exposed to misinformation that the tool carried was a hammer, the misinformation object is gist consistent (i.e., also a tool), and the participant may acquiesce to the suggested object. Thus, the gist trace for "tool" is processed and stored independently from the verbatim trace "screwdriver".

Younger children have a greater propensity to process verbatim details because they take longer to build up gist traces, as suggested by both fuzzy-trace theory (Brainerd & Reyna, 1990; Brainerd & Reyna, 1993; Brainerd & Reyna, 1998; Brainerd & Reyna, 2004) and Goodman's schema-confirmation-deployment theory (Farrar & Goodman, 1992). Older children build up gist traces, or scripts, relatively more quickly, and sometimes at the expense of the details (Brainerd & Gordon, 1994). For example, older children are more likely than younger children to recognize that the activity always begins with something to be sat upon, but may not have deeply processed each instantiation. In people of all ages, however, verbatim traces decay much more quickly than do gist traces (Brainerd & Reyna, 2004).

Source-Monitoring Framework. The source-monitoring framework (Johnson et al., 1993) while not a script-based theory, is relevant to an understanding of children's confusion concerning the source of details when recalling an instance of a repeated event.

Both fuzzy-trace theory and the source-monitoring framework provide an explanation as to why children may confuse details across multiple occurrences, when interviewed after a short delay.

The source-monitoring framework explains children's confusion of details across multiple occurrences as incorrect source attributions (Johnson et al., 1993). Therefore,

while fuzzy-trace theory posits that source is encoded, the source-monitoring framework suggests that source is attributed at recall.

The source-monitoring framework (Johnson et al., 1993) also provides an explanation for children's difficulty in recalling specific details of repeated events. Source-monitoring refers to the ability to consider the origins of a memory (Johnson et al., 1993). This task is made more difficult when sources are highly similar, such as with instances of a repeated event (e.g., Powell & Thompson, 1996), or when information is delivered by two similar speakers (Lindsay, Johnson, & Kwon, 1991). The identification of source, according to the source-monitoring framework, is a decision made at retrieval, rather than having been encoded, based on the amount and quality of perceptual information and cognitive operations information. The more perceptual detail a memory has (e.g., remembering how something smelled), the more likely it is to be a memory of an experienced event (Johnson & Raye, 1981). Memories that can be characterized as containing more thoughts and reflections, and less vivid perceptions, are more likely to be memories of imagined events (Johnson & Raye, 1981). Many source distinctions are not quite so straightforward, however. For example, when trying to determine which of two perceptually similar speakers mentioned a target detail, even adults do not perform at ceiling, especially when the speakers have discussed semantically similar content as well (Lindsay et al., 1991: experiment 1). Clearly a source decision is much more difficult under these circumstances, because the two sources are extremely similar in many respects.

Younger children have more difficulty than older children and adults in spontaneously making these types of distinctions, and have been shown to confuse

memories for imagining what someone has done versus having actually seen someone doing something (Lindsay et al., 1991). To complicate matters further, children who attempt to discriminate one in a series of repeated events may have equivalent amounts of perceptual information in their memories for each event.

Several studies have demonstrated improvements in source-monitoring ability somewhere between the ages of 3- to 8-years old (see Roberts, 2002, for a review of the development and theories of source-monitoring). For example, Drummey and Newcombe (2002) found dramatic increases in source memory for factual information given by two different speakers between the ages of 4- and 6-years old. The 4-year olds appeared to attribute much of the facts to extra-event sources, while the older children performed overall better but made between-speaker source errors.

In summary, all three theories suggest that children (and adults) should have difficulty recalling the exact details of one instance of a repeated event, and that younger children will struggle more than will older children. Script theory posits that script-consistent details will be absorbed into the script, such that when recalling the script, confusions across occurrences will be common. Since older children build scripts more quickly, they are better equipped than younger children to notice deviations. Fuzzy-trace theory suggests that when gist traces are activated, any gist-consistent detail could plausibly be recalled, however, if a verbatim trace is activated, and is still intact, memory for one instance of a repeated event should be good. Verbatim traces decay more quickly in younger children than in older children. Finally, the source-monitoring framework predicts confusions across occurrences because of the perceptual similarity between events. An assumption of this framework is that source is attributed through a decision-

making process at recall, and older children are more cognitively able to make these decisions than are younger children.

Variable versus Fixed Details

Repeated experience can have both positive and negative effects on children's memories for the events. If items that vary across occurrences are consistent with the overall script, they may become confused across occurrences, as predicted by fuzzy-trace theory (Brainerd & Reyna, 2004), and demonstrated in several experimental studies (e.g., McNichol, Shute, & Tucker, 1999; Roberts & Powell, 2006a). According to Graesser, Gordon and Sawyer's (1979) schema pointer plus tag model, items that vary, but are script-consistent, may not even be discriminated as being unique to the occurrence.

These studies have also, however, demonstrated the benefits of repeated experience on children's memory for details that never vary across occurrences ("fixed details") (Connolly & Lindsay, 2001; McNichol et al., 1999; Roberts & Powell, 2006a). Although children may be confused about items that are sometimes present and sometimes absent, repeated experience does appear to provide children with strong memory traces of details and events that are the same each time.

Memory for Single Experiences

Children typically have good, enduring memories for single-experience events (e.g., Fivush et al., 1984). Even children as young as 3-years old evidenced some recall of a laboratory-created novel event after 8-weeks (and no intervening interviews). Children 4 ½ years of age in the same study recalled significantly more general event details (main phases of the event), and objects (e.g., "a shirt") than the younger children, and only slightly more activity details (e.g., "hammering," "sewing") (Jones, Swift, & Johnson,

1988). Other research has also demonstrated that older children have more accurate memories for single-experience events than younger children (e.g., Powell et al., 1999).

Whether children are less suggestible about memories for unique experiences than children with memories for repeated events remains a debate in the literature. Some researchers have found that children who have engaged in repeated events were not more suggestible than children who took part in a novel event (Powell & Roberts, 2002), while other researchers have demonstrated that children who experienced multiple events were more suggestible (Connolly & Lindsay, 2001), using different methodologies in the interview. Notably, Powell and Roberts (2002) used cued recall questions rather than yes/no questions which children can plausibly answer without cognitive effort.

The type of detail studied also affects whether children with multiple experiences appear more or less suggestible. As children with repeated experiences have stronger memory traces for the details that never vary, this makes them more resistant to suggestions about those details than children who have experienced the event just one time (Powell et al., 1999). For details that vary, however, children with repeated experience sometimes make more source-errors than single-event children because the former remember multiple variations of a detail across multiple occurrences, especially when it is script-consistent (Roberts & Powell, 2006a). Additionally, the timing between the event(s) and the interview, of when potentially contaminating information might be presented, affects the picture of children's suggestibility both when they have had single or multiple experiences (see Roberts & Powell, 2007).

Source-confusions increase for children and adults when sources are highly similar (Johnson et al., 1993), such as with instances of repeated events. Children with

only one experience are likely to correctly attribute the detail to that occurrence if it has not been forgotten.

The more unique the single-event experience is, the less likely children will have developed a script about it. Scripts do help people organize memories (Abelson, 1981), thus it is likely that children with unique experiences tend to remember the most salient details. Scripts can be likened to gist traces, however, which, according to fuzzy-trace theory, are associated with greater suggestibility, when suggestions are consistent with the gist (e.g., Brainerd & Reyna, 2004). Verbatim traces, on the other hand, decay quickly but preserve the surface structure of the event. As mentioned previously, younger children also have a greater propensity to encode verbatim traces because it takes them longer to build up gist traces (Brainerd & Gordon, 1994). This idea is consistent with script theory in that younger children require more exemplars to build a script (Farrar & Goodman, 1992). Thus children with only one experience may have a stronger verbatim trace than gist trace.

Current Study

The bulk of the research described here has been concerned with adding to theories of children's cognitive development. These studies have engaged children in pleasant activities such as creative movement workshops (Hudson, 1990), and listening to stories and doing puzzles (Roberts & Powell, 2006a). These are well controlled studies in which the investigators have created the events and can compare the details mentioned by the child to those that actually occurred in the event to determine accuracy. Such research is imperative to an understanding of children's memories for repeated events. Results of

this work can be shared with non-academics who are involved in interviewing children, in order to further improve the quality of investigations with children.

While several theories make predictions about children's abilities to describe specific instances of a repeated event (e.g., schema theory models, schema-confirmation-deployment model, as reviewed in Farrar & Boyer-Pennington, 1999), none of these theories has directly examined ways of reducing errors in reports of a single instance of a repeated event. More specifically, none has compared the effects of practice in describing a repeated event, either using a script (e.g., what usually happens at swimming lessons), or by attempting to elicit the details associated with a specific instance. Such practice in describing an event from daily life is a technique that could potentially be used in the field if it can be demonstrated to improve children's narratives for an instance of a target repeated event.

Fuzzy trace theory (Brainerd & Reyna, 2004) suggests that incident-specific information ("verbatim traces") is processed separately from, and in parallel to, schematic information ("gist traces"). Although the theory does not make a direct prediction with respect to practice, it does posit that verbatim and gist retrieval are dissociated processes, and therefore, engaging in verbatim recall for one memory set (repeated event from child's daily life) may encourage continued verbatim recall for another memory set (target activities), although the theory does not explicitly make this prediction. Additionally, if a child has just engaged in gist recall for an autobiographical memory, (s)he is likely to continue in that mind-set for a target memory. From the perspective of script theory, incident-specific practice may help children describe an unrelated repeated event with more precision because this type of practice encourages

children to focus on episodic details of specific instances, rather than to rely on their scripts. Additionally, as in source-monitoring training studies where different sources are highlighted to children in training, incident-specific practice should demonstrate to children that there are different occurrences (sources) to consider when talking about a repeated event.

The current research proposes to contrast the effects of practice in describing a repeated event from a child's daily life, using general terminology (e.g., what happens at swimming lessons) with the effects of practice in describing a specific instance of a repeated event (e.g., what happened at swimming lessons on a particular occasion). A group of control children was given an equal amount of practice, but was asked to describe a single-experience event. Following the practice phase, children in the current study were questioned about target activities that had occurred previously in the lab (the "Laurier Activities").

It is known that practice in conversing with an interviewer about a recent event does improve the quantity of information given by children in forensic interviews (Orbach et al., 2000). Additionally, in an analogue study, Roberts, Lamb and Sternberg (2004) found that children whose rapport phase consisted of open-ended prompts generally provided more accurate accounts of a target event than children whose rapport phase consisted of directed questions (e.g., "How old are you?"). We do not currently know, however, if the level of specificity used by the child and interviewer during the practice phase will have an effect on the quantity, and quality, of information reported by the child, relative to other types of practice.

In the current study, 5-to 6- and 7- to 8-year old children practiced describing either two specific instances of a repeated event from their daily lives (**incident-specific practice**), or practiced what usually happens during an event they routinely experience (**general practice**). The control group described a novel event (**control practice**).

Incident specific practice may encourage the child to engage in thinking about the *differences* among multiple instances. Thus, in the incident-specific practice condition, children were asked to describe two unique occurrences in order to implicitly heighten awareness that some details of the two occurrences may differ. The interviewer also made this point more explicit by starting the practice phase with the following sentence: "some of the things we do when we do X are always the same, and some things are different." In order to even choose one instance to describe, the child must consider what was unique or relevant about a particular time. On the other hand, general practice may encourage the child to consider the *similarities* among multiple instances, and thus engage in a discussion of what "usually happens," or what is "typical." Although the source-monitoring framework does not make a direct prediction about the transfer of this type of practice, source-monitoring training studies to date have been grounded upon the notion that source training (e.g., practice in discrimination) using one set of event materials can be transferred to another set of materials (e.g., Poole & Lindsay, 2001; Roberts & Cameron, 2006; Thierry & Spence, 2002).

Following the practice phase, children were asked to describe "the time they remember best," when discussing the Laurier activities because research has shown that children have difficulty using temporal terms such as "yesterday," or "the first time" as retrieval cues (Hudson & Nelson, 1986). Specific cues tend to be even more useful for

retrieval, such as "What happens when you play with the cat?" (Farrar & Goodman, 1992). While we initially did not use such explicit prompts in the current study, it was hoped that children would be able to identify an instantiation of a variable detail on their own (e.g., "the time we wore *jellybean* badges"), in which case the interviewer requested more information based on the specific label provided by the child (e.g., "Tell me more about the time with the jellybean badges").

In order to report one specific instance of a repeated event, which contains both details that remain the same, and details that vary, children must first select one of these instances from memory, perhaps the time they remember the best. In most cases, this process began with the identification of a salient detail which could be used to label the instance. The label is only effective, however, if it is unique to one occurrence. Research has shown that children between the ages of 5- to 8-years old have some difficulty in choosing unique labels (Roberts & Powell, 2003). Once a label has been chosen, and the event script activated, the child must engage in a decision-making process about which details were present at that occurrence. In some cases, the source may have been encoded along with the detail, in the verbatim trace. If the verbatim trace is still intact, the detail may be matched to the correct source. If the gist trace is stronger than the verbatim trace, however, any gist-consistent instantiation of the detail may be retrieved (Brainerd & Reyna, 2004). If the verbatim trace has faded, source can still be correctly attributed to the detail through source-monitoring decisions. For example, the most recently experienced details often have the strongest perceptual features associated with them (Johnson et al., 1993), and therefore if a child has chosen to describe the *last time* something happened, source-monitoring decisions might be made effectively. It is clear

that describing one instance of a repeated event is not an easy task. The task is made even more difficult in real world situations such as forensic investigations, when the interviewer does not have access to all of the information and may not know at the outset whether the case involves repeated abuse or not. Such difficulties encountered in the field highlight the importance for academics to continue their challenging work in developing theoretically guided techniques that can be easily incorporated by non-academic professionals to assist children in delivering narratives of specific instances of repeated events that contain as much forensically relevant information as they can provide, without subsequent decreases in accuracy.

Hypotheses

Practice Phase. Manipulation checks are possible by analysing the quality of reports in the practice phase. The first hypothesis was that children in the specific practice condition were expected to use more episodic language in the practice phase, whereas children in the general practice condition were expected to use more scripted language. Script theory researchers (e.g., Fivush et al., 1984) have found that children do in fact respond in kind to the type of language used by an interviewer (i.e., if an interviewer asks what "happened," even very young children are likely to respond with equally specific, past-tense language). Children in the control condition were expected to use episodic language in practice because they are describing a single-experience event from the past. Additionally, older children in all conditions were expected to use more episodic language than were younger children.

Substantive Phase: Disclosure of Multiple Instances. Having practiced describing two specific instances of a repeated event in their daily lives, the second hypothesis was

that children in the incident-specific practice condition were expected to spontaneously disclose, earlier in the interview than the other two conditions, that the Laurier Activities happened more than once. Although none of the theories make explicit predictions with respect to this point, source-monitoring training studies are based upon the assumption that the skills acquired in practice will directly transfer to the interview (Poole & Lindsay, 2001; Thierry & Spence, 2002). To extend this idea, children in the specific practice condition should recognize that the Laurier Activities are another repeated event, and that the interviewer will expect them to talk about different times that they engaged in that event. Practice for the general condition children should encourage a script-like dialogue, where there is no emphasis placed on recalling separate instances. Finally, the control condition, having just described a unique event, may also fail to realize that the interviewer might be interested in knowing that more than one instance of the Laurier Activities occurred. It was also expected that older children would disclose earlier than younger children, because their narratives tend to contain more temporal markers (e.g., the first day) than the narratives of younger children.

Substantive Phase: Language Use. Children in all conditions were expected to use primarily episodic language, because research has shown that children with repeated experience interviewed according to the NICHD protocol largely tended to use episodic language to describe abuse, even when interviewer prompts were not always episodic (Elischberger, Pipe, Orbach, & Lamb, 2005, May). The NICHD protocol encourages episodic recall, and thus would be unlikely to contain a high number of generic interviewer utterances. Nevertheless, Elischberger and colleagues (2005, May) found that 25 of 36 interviews contained some child-provided generic information. When

interviewer utterances from that sample were examined, it was found that the average proportion of generic interviewer prompts was .39. Yet, in contrast, the average proportion of children's utterances that were generic was .08. Thus, while all children in the current study were expected to use a majority of episodic language, the third hypothesis was that children in the specific and control conditions would demonstrate the highest ratio of episodic to scripted language, while children in the general practice condition would have the lowest ratio. Children in the general practice condition were expected to provide more script-like narratives of the Laurier Activities, even if questioned about a specific instance of those activities, than children in the other two conditions.

While fuzzy-trace theory posits that gist and verbatim traces are stored independently, and therefore can potentially be retrieved independently given the proper support, script theories suggest that details absorbed into the script are not likely to be later accessible. If our data were to show that children in the incident-specific practice condition used more specific language than children in the general practice group, but still showed poor discrimination across occurrences, we would find support for script theories. On the other hand, if children in the incident-specific group used more specific language *and* showed better discrimination across occurrences, the data would be more in support of fuzzy-trace theory, suggesting that an intact verbatim trace was activated. The combination of more specific language and greater accuracy would also be a prediction of the source-monitoring framework, because to choose one instance to describe requires that a source decision was made.

The fourth hypothesis predicted that older children, and children in the specific practice condition would make more temporal references (e.g., "on Monday"), and more discriminatory references (e.g., "she brought a penguin *only once*") in their narratives of the Laurier Activities, than younger children, and children in the general and control conditions. Children in the general condition were expected to use more generic references (e.g., "she *always* brought her walrus") than other conditions.

Substantive Phase: Recall. Hypothesis 5 predicted age differences in the amount of detail recalled. Both field and analog research has shown that older children tend to report more details than do younger children (e.g., Farrar & Goodman, 1992; Lamb, Sternberg, Orbach, Esplin, Stewart & Mitchell, 2003). Condition differences were not hypothesized, and the analyses are exploratory, because it is possible that incident-specific practice would encourage children to recall as much as they possibly can about instances of the Laurier Activities, or alternately that the practice would make them more cautious about reporting details when they are unsure of the source, and thus they would report less than other conditions.

A main effect of detail was expected, in that fixed details should be reported with greater frequency than variable details, by all children, because they are present in every occurrence (hypothesis 6). Their memorability should be high, and children should not be confused about their source. When considered in the context of reporting details from one specific occurrence, fixed details should be reported with the highest accuracy, because they occur in every one of the four sessions.

Substantive Phase: Difference References. Hypothesis 7 predicted that children in the specific condition would recognize and mention more differences across sessions of

the Laurier Activities than children in other conditions, because that is another skill that they could have acquired through practice in describing two different instances of a repeated event. An interaction was also expected, in that 5- to 6-year old children in the specific condition may refer to more differences than other 5- to 6-year olds, while condition differences may not be observed among the older children, as older children naturally tend to notice deviations in routine events with greater frequency than younger children (Bauer, 2006).

Substantive Phase: Labels. In order to determine children's level of source-monitoring for each narrative (i.e., the proportion of details they correctly attribute to the session being described), the child and interviewer collectively must assign a "label" for the occurrence. Hypothesis 8 predicted that older children would spontaneously generate more of their own labels, whereas younger children were expected to have more interviewer-provided labels (chosen on the basis of something the child has already mentioned). Regardless of who provided the label, children in the specific condition were expected to have more unique labels than children in the other conditions, due to heightened awareness of differences across sessions (hypothesis 9).

Substantive Phase: Source Accuracy. Age differences were not predicted for internal source confusions (i.e., confusing details across different sessions of the Laurier Activities). Because the older children will likely have more gist-consistent instantiations in memory, they may have difficulty choosing the correct instantiation for the occurrence they are describing (Brainerd & Reyna, 2004). Although their source-monitoring skills should be better than those of the younger children, the four sessions will be highly similar semantically and perceptually, thus making source decisions difficult. On the

other hand, younger children may have fewer instantiations in memory, but will have more difficulty with source decisions (Johnson et al., 1993). In general, however, script theories, fuzzy-trace theory and the source-monitoring framework all suggest slightly better performance for the older children versus the younger children. With respect to the different types of details, the prediction of hypothesis 10 was that children of both age groups would report Hi/Lo details with more accuracy than variable details. Hi/Lo details are the same for three of the four occurrences ("Hi"), and change in one of the occurrences ("Lo"), making both the high and low frequency instantiations memorable.

If incident-specific practice indeed promotes more episodic language in children's reports of the Laurier Activities, it could be hypothesized according to fuzzy-trace theory that children in the specific condition would activate the verbatim trace associated with one instance of the Laurier Activities and thus have a higher source-accuracy score than children in other conditions. On the other hand, script-theories would predict only age differences, and not differences across conditions in source-monitoring, because script theories do not incorporate dual-process models of memory; only the script trace would persist, with any script-consistent instantiation plausible for a given occurrence.

In accordance with hypothesis 11, the younger children's narratives were expected to include more *external* intrusions (i.e., unrelated to the Laurier Activities) than those of the older children because they will not have developed a script of the activities as quickly as the older children (Farrar & Goodman, 1992) and their verbatim traces for each session will have decayed more quickly than those of the older children (Brainerd & Reyna, 2004). Thus, when attempting to describe an instance of the Laurier Activities, the older children may have a pool of instantiations to choose from in memory, whereas

the younger children may have fewer of these details to choose from and may simply provide information that is more recent in memory, such as details and events they have experienced the day before in school.

The main design of the study was a 2(age group: 5-6 years old, 7-8 years old) x 3(practice condition: specific, general, control). There were a number of dependent variables, such as the proportion accuracy of details provided. The influence of age and practice condition on these dependent variables was assessed through a series of analyses of variance (ANOVAs), or multivariate analyses of variance (MANOVAs) where appropriate.

Method

Participants

Initially, 157 children were recruited to take part in the research. The final sample consisted of 118 children. Of the 39 children who were excluded, 15 were lost due to parental time constraints, 10 discontinued their participation as the weather improved (sessions were run during recesses), 4 were ill for the interview session, 3 were excluded due to interviewer error, another 3 were not invited back to sessions because of behavioural problems, 2 children were too young, 1 child was selectively mute, and 1 child declined participation. Because children were assigned to their condition at various time points during their participation (i.e., when the researchers were able to obtain an event from the parent), differences among children who were lost were not compared. Many had not yet been assigned to any condition.

Fifty-eight 5- to 6-year olds and 60 7- to 8-year olds recruited from Waterloo Region and surrounding area took part in the current project. The sample consisted of 51

males and 67 females. The composition of males and females in the control, general, and specific conditions, respectively were; 19:19, 19:21 and 13:27. Effort was made to keep gender as balanced as possible across condition, however, due to the discrepancy in the specific condition, a variety of analyses were conducted to demonstrate that the males and females did not differ on variables of interest.

The overall sample mean for age was 83.51 months (6.96 years), range 60.10 - 108.13 months. In the control condition, the 5-6-year olds had a mean age of 72.56 months (*SD* = 8.33), range 60.93 – 83.07 months, and the 7-8-year olds had a mean age of 95.52 months (*SD* = 7.95), range 86.13 – 107.93 months. In the general condition, the 5-6 year olds had a mean age of 69.32 months (*SD* = 7.95), range 60.10 – 82.73 months, and the 7-8 year olds had a mean age of 94.90 months (*SD* = 6.81), range 84.43 – 106.53 months. In the specific condition the 5-6 year olds had a mean age of 72.57 months (*SD* = 5.56), range 64.10 – 80.47 months, and the 7-8 year olds had a mean age of 96.17 months (*SD* = 7.53), range 86.43 – 108.13 months. An ANOVA conducted on age in months, with condition as the independent variable indicated no difference in age across conditions, $p = ns$, $\eta_p^2 = .007$.

The ethnic and socioeconomic composition of the sample was expected to represent the natural diversity of the region, but no effort was made to specifically recruit children from any particular background. All children demonstrated enough proficiency in English to comfortably have a conversation with the interviewer. Children were recruited from a local daycare and public school in Waterloo, ON, as well as a public school in Wilmot Township, ON, and through the Laurier Family Database, which

contains the names and contact information of local families who have expressed interest in the research.

Informed consent was sought from parents through a signed Informed Consent form, and assent was acquired from all children who participated. Parents who were recruited from the Laurier Family Database received $15 in compensation for their time, and children received a small toy (approximate value: $1) at each session. Compensation for all children recruited elsewhere included a donation of $50 per participating grade and an additional $3 for each participating child. At the requests of the school/daycare directors, participating children were not compensated with toys. All participants were treated in accordance with the Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans.

Design and Procedure

All of the children participated in four 30 min sessions of the "Laurier Activities" over a two week period, and then were interviewed in one of three conditions 5- to 7-days following their final session, lasting 45 min. The six female interviewers were naïve, in that they did not know with which specific props the child interacted, and were blind as to the hypotheses of the study. In the "Control" condition, children were asked to describe a recently experienced *novel* event, provided by their guardian. Novel events were unique one-time experiences, and the children were not to have had similar experiences. For example, one child's novel event was a trip to Canada's Wonderland, a theme park in Ontario, and this child had not been to similar theme parks in the past. In the "General Practice" condition, children were asked to describe what usually happens when they take part in a repeated event from their daily lives (e.g., weekly swimming

lessons). These events were recorded with the children's data, and attempts were made to control as much as possible the frequency and recency of these events. For example, parents were asked to provide a repeated event which conformed to the following rule as much as possible: "a pleasant event that your child engages in roughly one time per week, and has been doing so for at least a month, but no more than four months." The four month limit is intended to prevent parents from choosing an overly familiar event. In the "Incident-specific practice" session, children were asked to describe what happened the *time they remembered best*, as well as one additional time, when they engaged in a repeated event from their daily lives. Once the narrative for the time remembered best was exhausted, children were asked to describe one other time that they remembered well. Because the aim of incident-specific practice was to highlight to children the importance of recognizing that instances of repeated events can have different elements, we had to give children the opportunity to describe more than one instance (see Appendix A for Interview Practice Phase guidelines).

The amount of time that children spent describing a novel event, or either what "usually happens" during a repeated event, or what "happened" the two times best remembered, was held as constant as possible across all interview conditions. Interviewers were given a target time frame of five to seven min for the practice phase in all three conditions.

Following the practice session, all children were questioned in an open-ended, neutral, non-suggestive manner about the Laurier Activities (see Appendix B for the protocol interview). Interviewing instructions were taken from the NICHD Investigative Interview Protocol (Orbach et al., 2000). The substantive phase of this interview begins

by eliciting a disclosure from the child as non-suggestively as possible, using increasingly more direct questions until a disclosure is obtained. Because the current study is not concerned with methods of obtaining a disclosure, the substantive phase of our interview simply began with an invitation to discuss the Laurier Activities. The entire interview was broken down into four main phases: the Practice phase, and three substantive phases: Narrative One, Restart Narrative One, and Narrative Two.

Narrative One began with the scripted prompt in which the interviewer indicated that she was naïve about the event, and asked the child to tell her everything (s)he could remember about the Laurier Activities. It is important to note that the Laurier Activities were always referred to in the plural; at each event session and at the interview, to represent the series of activities that happened in each session. Thus, referring to the activities in the plural at the interview should not have conveyed to the children that the interviewer was aware of event frequency.

Narrative One continued until either the child made a statement that indicated the Laurier Activities happened more than once (e.g., "and, on the last day, her cloak was red."), or until the child indicated that (s)he could not remember any more about the Laurier Activities. If the child disclosed multiple incidents, the interviewer immediately asked the child to tell her about the "time you remember best" of the Laurier Activities. If the child provided a more vague disclosure (e.g., she usually did X," "she always had X." or she sometimes did X") the interviewer asked the child for further clarification, which almost always immediately elicited a clearer disclosure.

If the child did not disclose multiple incidents, (s)he was asked if the Laurier Activities happened more than once after recall appeared to have been exhausted, as per

the NICHD protocol. The NICHD protocol recommends that if possible, children should themselves disclose repeated events. The protocol guidelines recommend asking the child about multiple occurrences only after recall appears to be exhausted. Children in the current study who had to be asked about multiple occurrences did not have a Restart Narrative One.

The "time you remember best" prompt began the phase "Restart Narrative One." For both Narrative One and Restart Narrative One, open-ended questions were used to elicit the children's memories such as Invitations (e.g., "Tell me more") and other general questions (e.g., "What else can you tell me about [detail already mentioned by child]?"). Once the interviewer felt that the child had provided as much detail for Narrative One (or Restart Narrative One, as the case may be), the first narrative was given a label. Labels were chosen either by the child or the interviewer. A child's label was chosen when the child had used a clear label throughout his/her first narrative (e.g., "the first time," "the time I wore a leaf badge"). Many children, however, do not so directly label occurrences of repeated events (Roberts & Powell, 2003a) and so, in these cases in the present study, the interviewer was required to choose a label using one of three strategies. All children were asked, prior to generation of the label, if there was anything different about the time they had just talked about, from any of the other times. This question encouraged the child to identify a unique feature of the occurrence (although not all children were accurate in doing so), which could be used as the label. Some children in the current study struggled with this question, even if they had previously identified a difference across occurrences. For example, many children indicated that they wore a badge every day, and that they were always different. If the child failed to identify any differences

when directly asked, (s)he was reminded of his/her earlier statement and was asked, for example, what badge was worn in the time the child had just described. If both of these strategies failed for choosing a label, the interviewers were instructed to choose one of the first instantiations mentioned by the child. Thus, because the interviewers were blind as to which counterbalanced group the children were in, and because children were sometimes inaccurate about what was unique, or failed to provide unique details, not all labels uniquely identified one occurrence.

After a label was generated for the first narrative, the children were invited to talk about another time when they participated in the Laurier Activities. This prompt began Narrative Two, which was also labeled.

Materials

The composite of props and activities presented to the children across the four sessions was as closely matched as possible to those used in some previous research on children's memory for repeated events (Pearse, Powell, & Thomson, 2003; Powell et al., 1999; Powell & Thomson, 2003; Roberts & Powell, 2005, 2006a, 2006b) modelled on Powell and Thomson's (1996) original "Deakin activities." The to-be-remembered events were made up of 17 target details which took place within the context of several activities: physical exercise, listening to a story, doing a puzzle, relaxing, getting refreshed and getting a surprise (see Appendix C for instantiations and counterbalancing). While some of these activities may be familiar to children (e.g., doing a puzzle), the individual props created for these activities were designed to be novel to all of the children who participated. As well, the sequence of activities that occurred was designed

specifically for the Laurier Activities and children were not expected to have any pre-existing scripts for these sessions.

Seven of the 17 details varied each time ("*variable items*," for example, some children heard a story about twins, a boat, winter activities, and a dog who moves to the city, across the four sessions). Five of the 17 details were the same each time ("*fixed items*," for example, children in another version of the counterbalancing heard the story about the dog four times). The remaining five details varied on a High-Lo frequency schedule ("*Hi/Lo items*"). High frequency details were the same for three of the four sessions, and the Lo-frequency detail was the instantiation of the item presented at the remaining session (e.g., clown juggling puzzle at sessions 1, 2, and 4; clown bicycle puzzle at session 3). The item that served as the Lo-frequency detail was cycled through each of the four sessions, such that there was one Lo instantiation of a Hi/Lo detail in each of the four sessions.

One target item was variable in both counterbalanced conditions: the badge that the child wore at each session (jellybean, button, leaves, pink feathers). Because the badge was worn during the entire session, it was expected that children might recognize it both as a salient detail, and as one that varies each time, and thus may be able to use it as a session-identifier during their interviews (e.g., "the time we had jellybean badges"), as was done in the study by Roberts and Powell (2006b).

Coding

Children's video- and audio-taped interviews were transcribed and sanitized (all identifying information such as names and places was removed, or replaced with initials). Coding began with a check of the interviewer's utterances. Coders verified that each

phase of the interview was carried out in the proper order and that the appropriate invitations for each phase were given. The number of additional prompts (e.g., "Tell me more about _____") was tallied for each phase. The total number of English words spoken by a child during the practice and substantive phases was also tallied.

Practice Phase. Because it was not possible to entirely verify each child's account of the repeated event from their daily lives, especially for children who participated in local schools, the practice portion of the interview was not coded for accuracy. Instead, this phase was coded primarily for amount of details provided and scripted versus specific language, as well as a count of number of child words and interviewer prompts.

Details are considered words relevant to the child's narrative, and only unique details are counted. For example, the phrase "my mom put the lifejacket on me" would be coded as five unique details for "my" "mom" "put_on" "lifejacket" and "me." Any other references to mom putting a lifejacket on the child were not counted as unique details, however, if the child said "I took off my lifejacket," four unique details were counted. This method of coding details is similar to that used in other research conducted by the lab of the author (e.g., Evans & Roberts, 2007), as well as in other published research (e.g., Orbach et al., 2000).

Language coding was carried out for both the practice and substantive phases of the interview. Phrases containing present-tense language (e.g., "*there are* lots of other kids there" and "we *play* with pool toys") were coded as scripted language. Phrases containing past-tense language (e.g., "we *played* with a beach ball") were coded as episodic language. Additionally, terms like "usually," and "always," were coded in a

category called "Generic references" which parallels scripted language, whereas terms such as "that time," and "only one time," and any other utterances that indicate a distinction between occurrences (e.g., "sometimes we have _____, and other times we have _____") were coded as "discriminatory references," which are generally indicative of the child using episodic language. A category for temporal references was also created (e.g., "on Monday," "last week"). Finally, a category for "Future references" (e.g., "and next week we're going to use the diving board") also was created, but was not included in any analyses as it was infrequently used. A ratio of episodic to scripted language was calculated because it was apparent that as the number of words increased, the figures in these categories increased as well. Generic, discriminatory and temporal utterances, although used by most children, were small in number, and thus proportions were not calculated.

In order to determine the amount each child talked in practice, the total number of English words produced by the child was counted. Interviewer prompts were also tallied. A prompt included any request by the interviewer for more information (e.g., "tell me more about X"), and did not include statements of interest or facilitators (e.g., "okay," "mm-hmm," "that sounds like fun").

Substantive Phase. The coding sections and rules described here were used to code the interviews from a sample of Australian children who took part in similar activities (Deakin Activities) (Roberts & Powell, 2006b). In the event that the authors may wish to make some comparisons between the samples, the same coding procedures were used.

Firstly, the phase in which the child disclosed multiple incidents (or had to be asked), was noted, along with the number of interviewer prompts in the substantive phase prior to it being evident that the Laurier Activities happened more than one time.

The substantive portion of the child's transcript was divided into "first narrative" (N1) and "second narrative" (N2) and was coded for amount and quality of information. Each of the 17 details, and each instantiation of those details that the child mentioned was recorded, along with its frequency (fixed, Hi/Lo, variable) and for instantiations, the session(s) in which it occurred (1, 2, 3, or 4) was also noted. Raw totals and proportions (i.e., proportion of instantiations reported that were fixed, Hi/Lo, and variable) were calculated for number of items and instantiations for each narrative.

External intrusions (i.e., indicating that there was a flower badge, when no such badge was present in any session) were also counted, and were coded as fixed, Hi/Lo or variable if they represented an Item-level detail present in the Laurier Activities. For example, "badge" was present in the activities as a variable detail, and "flower badge" was coded as a variable external intrusion. External intrusions that did not appear to represent any of the items in the Laurier Activities were coded as other (e.g., "we played Bingo").

Occurrence coding consisted of scoring each child's transcript based on the number of times the children spontaneously mention differences (e.g., "sometimes there was _____, other times there was _____"). These utterances were scored as accurate or inaccurate, for fixed, Hi/Lo and variable details separately.

The labels for the first and second occurrence were coded as to whether or not they uniquely identified a single occurrence of the Laurier Activities (however, for clarity

and brevity, analyses concerning labels, and source-accuracy for details associated with that label, will only be presented for the narrative one label in the current paper). For example, for Group 1, the colour of the cape is a variable detail and thus would uniquely identify one occurrence. The clown riding a bicycle puzzle is a Lo version of the Hi/Lo puzzle detail for the same group, and would also uniquely identify a single occurrence. Labels were either spontaneously identified by the child (e.g., "the time we had jellybean badges") or were agreed upon by the child and the interviewer (e.g., if the child was describing the time remembered best, and mentioned that there was a clown on a unicycle puzzle, but the puzzles were different in the other times, the interviewer would suggest labelling the recalled instance "the time with the unicycle puzzle"). After identifying and coding the effectiveness of the label for the described occurrence, all of the subsequent details mentioned by the child for that occurrence were recoded as to whether they were actually from the same occurrence as the label detail, or whether they were intrusions from other occurrences. Children received an "integrity score" for their narrative one, which was calculated by dividing the total number of details mentioned by the number of details that were actually present in the occurrence being described. This calculation was performed only on Hi/Lo and variable details because fixed details were always accurate by nature, and thus a child reporting only fixed details would appear to be 100% accurate.

While describing an instance of the Laurier Activities, some children mentioned that there were other instantiations that were present in a different occurrence of the Laurier Activities, and indicated the other occurrence (e.g., when describing the "first time," a child mentioned that the cloak was green and added "but the last time it was red."). These details were coded as "extras," and the occurrence mentioned by the child

was recorded (e.g., red cloak = occurrence 4). These details were later checked for accuracy and separate variables were created to represent raw totals and proportions of accuracy for *extra* Hi/Lo and variable details.

Finally, children's transcripts were coded for scripted versus specific language, as described in the practice phase coding. It was expected that children who were given practice in describing specific instances would use more specific language in their narratives of the Laurier Activities than would children in the general practice condition. Total number of child words was also tallied for the substantive phase.

Reliability. Coders were all senior undergraduate students, and were trained by the author in coding procedures on approximately 10% of the transcripts. After training, an additional 15% of transcripts (18) were double-coded for reliability purposes. Percent agreement (number of agreements / number of agreements + disagreements) was used to assess reliability, except in the case of Language Coding, where correlations were performed due to the high number of coding categories. Percent agreement and correlations were greater than .90 for all coding. Reliability was high, likely due to straightforward coding schemes where variables of interest were "counted" and rarely involved subjective decision-making. An additional 6 transcripts were double-coded after coders had completed approximately two-thirds of the transcripts, to ensure that coding remained consistent. Percent agreement was consistent with earlier reliability assessments. Additionally, all data were double-entered by the principal investigator and coders, to ensure accuracy. When differences were observed, the transcripts and coding sheets were checked, and the correct value was entered into the data file.

Results

Practice Phase

In three-between subjects conditions, children received either: **incident-specific** practice, in which they were encouraged to report details associated with two instances of a repeated autobiographical event; **general** practice, in which they were encouraged to describe what generally happens when they engage in a repeated autobiographical event; or **control** practice, in which children were asked to talk about a novel, single-experience event.

Overall, children averaged 442.63 words (SD = 260.88) and 198.51 details (SD = 125.81) in their practice phases. Number of words, and number of details in practice served as the dependent variables in a 2(age group) by 3(condition) MANOVA. A main effect of age group emerged $F(2, 111) = 19.55, p < .001, \eta_p^2 = .26$. Condition was marginally significant, $F(4, 224) = 2.14, p = .077, \eta_p^2 = .04$, and the interaction was not significant, $F(4, 224) < 1.00, p = ns, \eta_p^2 < .02$. Follow-up univariate tests indicated that for number of words, older children spoke more (M = 541.87, SD = 271.35) than did younger children (M = 339.97, SD = 205.89), $F(1, 112) = 21.56, p < .001, \eta_p^2 = .16$. Additionally, there was a main effect of condition, $F(2, 112) = 4.41, p < .05, \eta_p^2 = .07$. Post hoc tests ($LSD, p < .05$) indicated that children in the specific (M = 504.10, SD = 295.57) and control (M = 471.74, SD = 256.68) conditions provided more words than did children in the general condition (M = 353.50, SD = 203.62), and did not differ from each other. The interaction component was not significant, $F(2, 118) = .56, p = ns, \eta_p^2 = .01$. For number of details, older children (M = 255.33, SD = 129.96) provided more details than did younger children (M = 139.72, SD = 89.96), $F(1, 112) = 32.39, p < .001, \eta_p^2 =$ .22, and children in the specific (M = 224.73, SD = 142.23) and control conditions (M = 212.74, SD = 124.32) provided more details than did children in the general condition (M

= 158.78, *SD* = 100.36), $F(2, 112) = 3.92$, $p < .05$, $\eta_p^2 = .07$, as indicated by post hoc analysis (*LSD*, $p < .05$). The interaction was not significant, $F(2, 112) = .21$, $p = ns$, $\eta_p^2 < .01$.

Number of interviewer prompts given in practice was examined in a 2(age group) by 3(condition) ANOVA. Main effects of age and condition emerged. Younger children (*M* = 16.21, *SD* = 7.99) received significantly more prompts than did older children (*M* = 12.73, *SD* = 6.60), $F(1, 112) = 7.24$, $p < .05$, $\eta_p^2 = .06$. Children in the specific condition (*M* = 17.67, *SD* = 8.44) received more prompts than did children in the general (*M* = 12.45, *SD* = 5.44) and control conditions (*M* = 13.13, *SD* = 7.24), $F(2, 112) = 6.51$, $p < .01$, $\eta_p^2 = .10$, and post hoc analysis (*LSD*, $p < .05$) revealed that only children in the specific condition differed from other groups. Since only children in the specific condition practiced talking about *two instances* of an event, we were interested in knowing whether the increased amount of prompts for the specific condition was due to the second practice phase. A similar ANOVA with only prompts in *practice one* (or just "practice" for the general and control group) was conducted. Only the main effect of age was significant, with younger children (*M* = 13.71, *SD* = 6.33) receiving more prompts than older children (*M* = 10.98, *SD* = 5.14), $F(1, 112) = 6.92$, $p < .05$, $\eta_p^2 = .06$. The main effect of condition, and the interaction were not significant, $F\text{s} \leq 1.36$, $p = ns$, $\eta_p^2\text{s} \leqslant .02$.

Finally, manipulation checks were conducted to determine whether children (and interviewers) in each practice condition were using the type of language that we had predicted in hypothesis 1 (i.e., episodic versus scripted). A ratio of episodic to scripted language was calculated, and used as the dependent variable in a 2(age group) by 3(condition) ANOVA. As expected, only the effect of condition was significant, *F*(2,

112) = 84.54, $p < .001$, $\eta_p^2 = .60$. Post hoc tests (*LSD*, $p < .05$) revealed that all conditions differed significantly from each other, with the control condition ($M = .86$, $SD = .19$) having the greatest ratio of episodic to scripted language, followed by the specific condition ($M = .73$, $SD = .28$), while the general condition ($M = .20$, $SD = .24$) had the lowest ratio. The main effect of age, and the interaction, were not significant, $Fs \leq .53$, p = ns, $\eta_p^2 s \leq .01$. Thus, contrary to expectation, children of both ages used similar ratios of episodic to scripted language. It is important to note that the inverse of this ratio represents the amount of scripted language used, and thus it can be inferred that children in the general condition had the highest ratio of scripted to episodic language.

A similar ratio for episodic to scripted language was calculated for the interviewer utterances, and entered into a 2(age group) by 2(condition) ANOVA. Again, only the effect of condition was significant, $F(2, 112) = 280.86$, $p < .01$, $\eta_p^2 = .83$. Post hoc analysis (*LSD*, $p < .05$) indicated that, as expected, interviewers in the specific ($M = .83$, $SD = .22$) and control conditions ($M = .87$, $SD = .20$) used a greater ratio of episodic to scripted language than did interviewers in the general condition ($M = .04$, $SD = .09$). The main effect of age, and the interaction component, were not significant, $Fs \leqslant .87$, $ps = ns$, $\eta_p^2 s \leqslant .02$.

Substantive Phase

Preliminary Analyses. Because gender was not balanced in the specific condition, the effects of this variable were assessed on all of the analyses that will follow in the Substantive portion of the results. Of 17 tests in which performance across genders was compared, only one produced a significant difference. Thus data were collapsed across gender for all analyses. Additionally, because the practice phase of the specific condition

appeared to be longer (in number of words, details, and interviewer prompts), number of interviewer prompts was used as a covariate in the analyses. The results of one ANOVA were affected by this variable, and the ANCOVA will be reported for that test only. Interviewer prompts in practice was used as the covariate, rather than number of child words or details, because it was the interviewers' prompts that encouraged words and details from the child.

Number of child words in the substantive phase was entered as a dependent variable into a 2(age group) by 3(condition) ANOVA. The only significant effect was of age, $F(1, 112) = 27.28$, $p < .001$, $\eta_p^2 = .20$, with older children ($M = 952.63$, $SD = 387.36$) using more words than younger children ($M = 588.71$, $SD = 362.34$). The effect of condition and the interaction were not significant, Fs $\leqslant .74$, $ps = ns$, $\eta_p^2 \leqslant .01$.

Also of note, several 2(age group) by 3(condition) ANOVAs had significant Levene's tests, indicating problems with homogeneity. Often it was the case that younger children evidenced greater variability than did older children. Standardizations and log transformations were applied to the variables of interest, however, findings did not change, and problems with homogeneity were not resolved. All reported dependent variables were also explored for outliers. No outliers of great concern were observed, and there was no systematic pattern, by age or condition. Additionally, no case (i.e. child) stood out as an outlier across multiple dependent variables.

Disclosure of multiple incidents. Hypothesis 2 predicted that children who practiced describing specific instances of repeated events would disclose earlier in the Substantive portion of the interview that the Laurier Activities happened more than once. A 3(condition) by 4(location of disclosure: immediate, breadth/depth, labelling, other) chi

square was conducted. More children in the specific condition than would be expected disclosed immediately, while fewer children than expected in the general condition did so. Additionally, more children than expected in the control condition disclosed in labelling (i.e., these children had to be asked by the interviewer if the LA happened more than once), and only one child in the specific condition had to be asked, χ^2 (6, N = 118) = 12.64, $p < .05$. When the file was split by age, and the same chi square analyses were run, it became evident that these effects were only true for the younger children, χ^2 (6, N = 58) = 13.51, $p < .05$, and not for older children, χ^2 (6, N = 60) = 3.86, $p = ns$ (see Table 1).

Another way to examine when children disclosed multiple incidents was to look at the number of prompts required by the interviewer in the substantive phase to elicit a disclosure of multiple incidents. This variable served as the dependent variable in a 2(age group) by 3(condition) ANCOVA, which included number of prompts in practice as a covariate. The two children who terminated early, and the two children who said the Laurier Activities only happened once, are excluded from this analysis. The variances between groups were not homogenous, with the younger children having greater variability, but transformation of the data did not resolve this issue. There was a main effect of age, in that older children required fewer prompts to disclose multiple incidents (M = 3.75, SD = 3.50) than did younger children (M = 8.80, SD = 6.61), F(1, 107) = 21.40, $p < .001$, $\eta_p^2 = .17$. Condition was significant, F(2, 107) = 3.48, $p < .05$, $\eta_p^2 = .06$. Planned t-tests indicated that children in the specific condition required marginally fewer prompts (M = 4.90, SD = 4.90) than did children in the control condition (M = 7.03, SD = 6.56), t(68.42) = 1.61, $p = .056$ (one-tailed). Children in the general condition required an

average of 6.54 prompts (*SD* = 5.69), and did not differ from either condition. No other differences were found.

Although the interaction component was not significant, $F(2, 112) = 1.42$, $p = ns$, $\eta_p^2 = .03$, it was noted that nearly all of the older children disclosed quite early and showed little variability, but that for the younger children, the mean number of prompts to disclosure appeared much lower for the specific condition relative to the other two. Planned *t*-tests run only on the younger children demonstrated similar effects as were obtained with the full sample. Younger children in the specific condition required an average of 6.63 prompts (*SD* = 5.91) while those in the control condition required an average of 10.78 prompts (*SD* = 6.94), $t(35) = 1.96$, $p = .029$ (one-tailed). Younger children in the general condition required an average of prompts 9.12 (*SD* = 6.64), and did not differ from either condition.

Language Use in Substantive Phase. Hypothesis 3, that children in the specific practice condition would use a greater ratio of episodic to scripted language in their substantive narratives than would children in the general practice condition, was supported. A 2(age group) by 3(condition) ANOVA was conducted, and only the main effect of condition was significant, $F(2, 112) = 14.73$, $p < .001$, $\eta_p^2 = .21$. Post hoc tests (*LSD*, $p < .05$) indicated that the specific (*M* = .88, *SD* = .14) and control (*M* = .88, *SD* = .16) conditions both used a greater ratio of episodic language in their narratives than did the general condition (*M* = .66, *SD* = .29). The greater variability observed in the general condition resulted in another significant Levene's test, which was not resolved through transformations of the data. The main effect of age was marginally significant, $F(1, 112) = 3.05$, $p = .084$, $\eta_p^2 = .03$, with younger children having a lower ratio of episodic to

scripted language (M = .77, SD = .25) than older children (M = .84, SD = .20). The interaction was not significant, $F(2, 112) = 2.30$, $p = .105$, $\eta_p^2 = .04$. Upon examination of the means, however, it was noted that the 5- to 6-year old children in the general condition (M = .59, SD = .30) had a noticeably lower mean than was observed in the other cells (See Table 2 for all means). Given the exploratory nature of the research, condition differences were analyzed for each age group separately. Condition was significant for both age groups, Fs $\geq$ 4.98, $ps \leq$,01, and it remained the case that children in the incident-specific and control practice groups had a higher ratio of episodic to scripted language than the general practice condition, according to post hoc tests (*LSD*, $p < .05$).

A series of 2(age group) by 3(condition) ANOVAs were run on the number of temporal, generic and discriminatory references provided by the children in the substantive phase. Hypothesis 4 predicted that older children, and children in the incident-specific practice condition, would use more temporal and discriminatory references than younger children, and all children in general and control conditions. Additionally, children in the general condition were expected to use more generic references. For all three variables, however, only the main effect of age was significant, Fs $\geq$ 7.93, ps < .01, η_p^2s $\geq$ .07. See Table 3 for all means.

Recall. To address hypothesis 5, that older children would recall overall more details than younger children, total number of *items* (i.e., a higher-order category present in every session, such as a *doing a puzzle* each time), and total number of *instantiations* (i.e., variations of the items, such as *the clown juggling puzzle,* and *the clown on a tightrope puzzle*) were entered into two separate ANOVAs with age group and practice

condition as the independent variables. Older children provided significantly more *items* ($M = 8.32$, $SD = 3.47$) than did younger children ($M = 6.12$, $SD = 3.40$), $F(1, 112) = 12.99$, $p < .001$, $\eta_p^2 = .10$. There was no main effect of condition, $F(2, 112) = .79$, $p = ns$, $\eta_p^2 = .01$, however, the interaction component was significant, $F(2, 112) = 4.46$, $p < .05$, $\eta_p^2 = .07$. Specifically, for the younger children there was an effect of condition, $F(2, 55) = 4.49$, $p < .05$, $\eta_p^2 = .14$, whereas for the older children there was not, $F(2, 57) = 1.01$, $p = ns$, $\eta_p^2 = .03$. Younger children in the specific condition ($M = 7.85$, $SD = 3.91$) provided significantly more items than did children in the general condition ($M = 5.05$, $SD = 2.84$), and control condition ($M = 5.39$, $SD = 2.68$), as indicated by post hoc analysis (*LSD*, $p < .05$).

With respect to *instantiations*, a main effect of age emerged, as expected. Older children ($M = 18.97$, $SD = 7.99$) provided significantly more instantiations than did younger children ($M = 10.40$, $SD = 6.70$), $F(1, 112) = 40.58$, $p < .001$, $\eta_p^2 = .27$. There was again no main effect of condition, $F(2, 112) = 1.35$, $p = ns$, $\eta_p^2 = .02$. The interaction was not significant, $F(2, 112) = 1.96$, $p = .146$, $\eta_p^2 = .03$, but due to the results of the previous analysis, each age group was separately subjected to an ANOVA to compare differences across conditions. Once again, condition differences were significant for younger children only, $F(2, 55) = 3.69$, $p < .05$, $\eta_p^2 = .12$. Post hoc tests (*LSD*, $p < .05$) indicated that younger children in the specific condition ($M = 13.15$, $SD = 7.33$) provided more instantiations than children in the general condition ($M = 7.65$, $SD = 5.58$), and neither group differed from the control condition ($M = 10.39$, $SD = 6.17$).

Hypothesis 6 predicted that fixed details (both items and instantiations) would be reported with greater frequency than would variable details, by children of both age

groups, regardless of condition, because they occur in every session. This hypothesis was not supported. A mixed 2(age group) by 3(condition) by 3(type of item: fixed, Hi/Lo, variable) ANOVA, the last factor within-subjects, was conducted. The test of sphericity was significant, and a Greenhouse-Geisser correction was applied. An effect of item-type was found, $F(1.71, 191.85) = 22.33, p < .001, \eta_p^2 = .17$. Follow-up paired-samples t-tests indicated that, unexpectedly, variable items ($M = 3.26$, $SD = 2.14$) were reported more frequently than were fixed items ($M = 2.92$, $SD = 1.91$), $t(117) = 4.33, p < .001$, and more than Hi/Lo items ($M = 1.86$, $SD = 1.38$), $p < .001$. No other within-subjects effects were found; all Fs ≤ 2.71, $p = ns$, η_p^2s $\leq .02$. The between-subjects main effect of age was significant, $F(1, 112) = 12.99, p < .001, \eta_p^2 = .10$; older children ($M = 2.77$, $SD = 1.16$) averaged significantly higher recall for items than did younger children ($M = 2.04$, $SD = 1.13$). The age by condition interaction was also significant, $F(2, 112) = 4.46, p < .05, \eta_p^2 = .07$. Separate ANOVAs for each age group revealed that the effect of condition was significant only for the younger children, $F(2, 55) = 4.49, p < .05, \eta_p^2 = .14$, not the older children, $F(2, 57) = 1.01$, $p =$ ns, $\eta_p^2 = .03$. For the younger children, post hoc analyses (LSD, $p < .05$) showed that children in the specific condition ($M = 2.62$, $SD = 1.30$) recalled overall more items than did children in the general ($M = 1.68$, $SD = .95$) and control conditions ($M = 1.80$, $SD = .89$), which did not differ from each other.

A mixed 2(age group) by 3(condition) by 2(type of instantiation: fixed, variable) ANOVA, the last factor within-subjects, was conducted. Again it was expected that fixed instantiations would be reported with greater frequency than would variable instantiations and this hypothesis was also not supported. Hi/Lo instantiations were not included in the analyses because this variable combines High and Lo frequency details together. A main

effect of instantiation type was observed, $F(1, 112) = 47.39, p < .001, \eta_p^2 = .30$, with variable details ($M = 6.63$, $SD = 4.62$) being reported more frequently than fixed ($M = 4.25$, $SD = 2.73$). An interaction between age and instantiation type was also observed, $F(1, 112) = 4.96, p < .05, \eta_p^2 = .04$. Younger children reported an average of 1.62 ($SD = 3.32$) more variable details than fixed, while older children reported an average of 3.12 ($SD = 4.09$) more variable details than fixed. A follow-up independent-samples *t*-test revealed these differences to vary significantly across age groups, $t(116) = -2.18, p < .05$. Finally, a main effect of age emerged, $F(1, 112) = 33.51, p < .001, \eta_p^2 = .23$. Older children ($M = 6.96$, $SD = 3.20$) had a higher average recall of instantiations than did younger children ($M = 3.86$, $SD = 2.57$). No other effects were significant, $Fs \leq 1.93$, $ps = ns$, $\eta_p^2 s \leq .03$.

Difference References. To test hypothesis 7 that children in the specific condition would be more likely than children in other conditions to make references to differences across sessions of the Laurier Activities, the total number of differences mentioned, the number of accurate difference references to variable instantiations, and the number of accurate difference references to Hi/Lo instantiations were entered into a 2(age group) by 3(condition) MANOVA. The main effect of age was significant, $F(3, 110) = 4.61, p < .01, \eta_p^2 = .11$, as well as the interaction, $F(6, 222) = 3.02, p < .01, \eta_p^2 = .08$. Condition was marginally significant, $F(6, 222) = 2.08, p = .057, \eta_p^2 = .05$.

Follow-up univariate analyses indicated that, for the total number of difference references, older children ($M = 2.40$, $SD = 1.44$) referred to differences significantly more than did younger children ($M = 1.45$, $SD = 1.58$), $F(1, 112) = 13.62, p < .001, \eta_p^2 = .11$. The effect of condition was significant, $F(2, 112) = 5.00, p < .01, \eta_p^2 = .08$. Post hoc

analyses (*LSD*, $p < .05$) indicated that children in the specific condition ($M = 2.50$, $SD = 1.70$) provided more difference references than did children in the control condition ($M = 1.66$, $SD = 1.32$), and general condition ($M = 1.63$, $SD = 1.56$), the latter two conditions not differing from each other. The main effects were subsumed by a significant interaction, $F(2, 112) = 4.31$, $p < .05$, $\eta_p^2 = .07$. Specifically, the effect of condition was significant only for the younger children, $F(2, 55) = 9.75$, $p < .001$, $\eta_p^2 = .26$. Post hoc analysis (*LSD*, $p < .05$) again revealed that younger children in the specific condition ($M = 2.55$, $SD = 1.80$) provided significantly more difference references than younger children in both the control ($M = .94$, $SD = 1.10$) and general conditions ($M = .80$, $SD = 1.15$). The effect of condition was not significant for the older children, $F(2, 57) < 1$, $p = ns$, $\eta_p^2 < .01$.

For accurate difference references to variable instantiations, effects of age and condition were found, Fs ≥ 4.90, ps $< .01$, η_p^2s $\geq .08$. Although the interaction component was not significant, $F(2, 112) = 1.67$, $p = ns$, $\eta_p^2 = .03$, the younger children in the specific condition had higher accuracy scores than did the younger children in the other conditions. For accurate difference references to Hi/Lo instantiations, the effect of age and the interaction were significant, Fs ≥ 6.26, ps $< .05$, η_p^2s $\geq .05$. The interaction component was explored in separate ANOVAs for each age group. Condition differences were significant for younger children, $F(2, 55) = 6.01$, $p < .001$, but not for older children, $F(2, 57) = 1.81$, $p = ns$. Post hoc analysis (*LSD*, $p < .05$) on the younger children's data revealed that 5- to 6-year olds in the specific condition provided more accurate difference references for Hi/Lo instantiations than children in the other two groups (see Table 4 for means). Reported scores are raw numbers, as inaccurate

references to variable and Hi/Lo instantiations being different across occurrences were extremely rare.

Labels. In order for the child and the interviewer to discuss the same occurrence of the Laurier Activities, a label for the occurrence was generated. For simplicity and succinctness, all following analyses in this paper will deal with the label generated for Narrative One only. Only eight children provided narratives for which labels were not generated, either due to early termination or saying that the Laurier Activities happened only once, or because the child only provided items and not instantiations. Seven of these children were 5- to 6-year olds (control = 3, general = 2, specific = 2) and one child was a 7- to 8-year old in the general condition. Thus, 110 labels were created, 54 (49.10 %) of which were variable (e.g., leaf badge), and 29 (26.40 %) of which were temporal (e.g., first time). Of the 54 variable labels, 17 were badges (31.48 %). Of the 29 temporal labels, 24 (82.76 %) referred to the first or last occurrence. There were no age or condition differences on the occurrence referred to by the labels, assessed by a series of chi squares.

Labels were spontaneously generated by the child, or chosen by the interviewer as needed. A series of chi squares was conducted to determine whether there were age or condition differences in who provided the label. Hypothesis 8 predicted that older children would be more likely to generate their own labels while younger children would be more likely to have interviewer-provided labels. A 2(label provider) by 3(condition) chi square, with file split by age, revealed no differences between observed and expected values for younger children χ^2 (2, $N = 51$) = 3.70, $p = ns$. For older children, however, those in the control condition required more interviewer-generated labels than expected,

whereas those in the specific condition generated marginally more of their own labels than would be expected, χ^2 (2, N = 59) = 5.54, p = .063. An alternate 2(label provider) by 2(age group), with file split by condition, revealed no differences between age groups in general and specific conditions, $\chi^2 \leq .51$, ps = *ns*. Older children in the control condition, however, required more interviewer-generated labels than would be expected, whereas the younger children in that condition generated more of their own labels than would be expected, χ^2 (1, N = 35) = 8.44, $p < .01$.

Not all labels were unique, because interviewers were blind as to which instantiations were present in each counterbalanced group, and thus did not know which instantiations were unique for each group, and because children themselves did not always choose unique labels. An independent samples *t*-test revealed that children with non-unique labels (n = 21) provided significantly fewer variable instantiations (M = 3.90, SD = 3.43) in their narratives than did children with unique labels (M = 7.83, SD = 4.34), $t(108) = -3.87$, $p < .001$.

Of the 110 labels, 89 (81 %) were unique. A 2(age group) by 2(unique, not unique) chi square revealed that younger children had more non-unique labels, and older children had more unique labels, χ^2 (1, N = 110) = 6.56, p = .01. Splitting the file by age and running a 3(condition) by 2(unique, not unique) chi square revealed a marginally significant effect for the younger children, χ^2 (2, N = 51) = 5.72, p = .057, but not older children χ^2 (2, N = 59) = 1.01, p = *ns*. Specifically, younger children in the general condition had more non-unique labels, and younger children in the specific condition had more unique labels, than would be expected. Partial support was found for hypothesis 9

as the prediction was that *all* children in the specific condition would have more unique labels than expected.

Source Accuracy. In order to determine whether children were accurate with respect to the instantiations they reported in their narrative, they had to provide a label (n = 110), and it had to be unique (i.e., refer to one occurrence only: a temporal label, a variable label, or a Lo instantiation of a Hi/Lo detail). In preliminary analysis of accuracy data, it was noted that some cells had low ns. A dummy variable was created to represent interviews in which a label was successfully generated and could be used to determine accuracy. A 2(successful, unsuccessful) by 2(age group) chi square, with file split by condition, was conducted. Age differences were non-significant in the control and specific conditions, χ^2s $\leqslant$ 1.76, $p = ns$. In the general condition, however, younger children were much less successful, and older children more successful, than would be expected in achieving a label that could be used to score accuracy, χ^2 (1, N = 40) = 9.23, $p < .01$. Less than half (n = 9) of the 5- to 6-year old children in the general condition met the criterion of providing a label, and having it be unique.

Children who did provide unique labels could be assessed for their source-accuracy (i.e., whether they retrieved details from the occurrence referred to in the label). Predictions were not made with respect to accuracy, although it was suggested from the theoretical frameworks reviewed that older children would be slightly more accurate than younger children. The children's integrity score for narrative one served as the dependent variable, in a 3(practice condition) by 2(age) ANOVA. Integrity scores were created by averaging children's proportion accurate scores for Hi/Lo and variable instantiations only, as source accuracy for fixed instantiations is 100%.

For narrative 1 integrity score, only a marginal main effect of age emerged $F(1, 74) = 3.50, p = .065, \eta_p^2 = .05$. Older children were slightly more accurate ($M = .61$, $SD = .27$) than were younger children ($M = .51$, $SD = .30$). Although no main effect of condition emerged; the control ($M = .59$, $SD = .29$) and specific ($M = .59$, $SD = .29$) had higher integrity scores than did the general condition ($M = .53$, $SD = .27$), $F(2, 74) = 1.11, p = ns, \eta_p^2 = .03$. The interaction was also non-significant, $F(2, 74) < 1.00, p = ns, \eta_p^2 = .01$.

Hypothesis 10 predicted that all children would report Hi/Lo details with more accuracy than variable details (i.e., attributing them to the correct occurrence). A 2(age group) by 3(condition) by 2(proportion accuracy: variable instantiations, Hi/Lo instantiations) mixed ANOVA was conducted. A main effect of proportion accuracy emerged, $F(1, 57) = 22.03, p < .001, \eta_p^2 = .28$. Children more accurately attributed Hi/Lo instantiations ($M = .74$, $SD = .31$) to occurrences than they did variable instantiations ($M = .46$, $SD = .40$). No other effects were observed, $Fs \leq 2.56$, $ps = ns$, $\eta_p^2 s \leq .04$.

When considering the "extra" instantiations provided by the children, there were not enough children in each condition who provided both unique labels, and extra Hi/Lo or variable instantiations from either narrative, to make analyses worthwhile. In addition, predictions were not made with respect to these variables.

Hypothesis 11 suggested that the younger children's narratives would contain more external intrusions than would those of the older children, regardless of condition. A 2(age group) by 3(condition) ANOVA was run on the number of external intrusions. No significant effects emerged, all $Fs < 1.84$, $ps = ns$, $\eta_p^2 s < .03$. The number of external intrusions provided by younger children ($M = .69$, $SD = 1.17$), however, was higher than

that of older children (M = .45, SD = .81). Overall, external intrusions were very low. The grand mean of external intrusions was .57 (SD = 1.01, range: 0 – 5). Twenty-three children provided one external intrusion and 79 of the children did not provide any.

Discussion

The goal of the current study was to determine whether a simple change to procedures already used by investigators in the field might be useful in eliciting more precise narratives of repeated events from children. Specifically, we were interested in determining whether practice in describing specific instances of a repeated event would benefit children's overall narratives.

If practice in describing specific instances of a repeated event from a child's daily life improves the quality and precision of that child's narrative for a target repeated event, such a technique could be easily employed by investigators in the field. It does not require that investigators have knowledge of the target event, and rather than being an additional procedure to what an investigator might already do, it simply provides a complement to an interview practice already in place.

Discussion of Overall Results

Across analyses, it was apparent that incident-specific practice had a greater effect on the reports of younger children than older children. The findings will be discussed first in terms of differentiating the behaviours of the incident-specific condition from the other two conditions. Secondly, findings related to each hypothesis will be reviewed.

Incident-Specific Practice. Most of the language used by the children in the incident-specific practice condition, both in the Practice and Substantive Phases can be characterized as episodic. Additionally, younger children in this condition disclosed

multiple incidents earlier, reported more items and instantiations (i.e., more forensically relevant details), mentioned more differences across occurrences, and more often provided narratives that led to the generation of unique labels, than their same-aged peers in the other conditions.

General Practice and Control. The main difference between these conditions is the type of language used both in the Practice and Substantive Phases. While children in the control condition both practiced describing an episodic event, and used episodic language to describe the Laurier Activities, children in the general condition practiced using scripts, and this type of language was more apparent in their descriptions of the Laurier Activities than for the other conditions. Additionally, while the children in the control condition required marginally more prompts to elicit a disclosure than incident-specific children, the general condition children did not differ from either group. On the other hand, younger children in the general condition provided narratives that were less successful in leading to the generation of a unique label, in contrast to their same-aged peers in the incident-specific condition. The two conditions were similar, however, in that younger children in these conditions were characterized by lower levels of recall and a lower frequency of references to differences across sessions.

Summary of Findings

Practice Phase. Children in all practice phases responded to the interviewer's language as was predicted by hypothesis 1; this finding is consistent with script-theory literature (e.g., Fivush et al., 1984), and thus it can be inferred that there were qualitative differences in the type of practice children received.

Although time limits and guidelines were in place to attempt to equalize the *amount* of practice children received across conditions, interviewers did use more prompts when talking to children in the specific condition. When analyses controlled for discussion of only one event, however, condition differences disappeared. Thus it was the discussion of a second instance of a repeated event that resulted in more prompts for the specific condition. The total number of prompts in practice was tested as a covariate in all main analyses and affected the results of only one analysis, suggesting that observed effects can likely be attributed to quality, and not amount, of practice. Additionally, number of child words and details did not differ between specific and control conditions, yet some analyses revealed differences between these conditions.

Substantive Phase: Preliminary Findings. In the substantive phases of their interviews, older children spoke more than did younger children, though no condition differences were observed. While it was expected that children in the specific condition might be encouraged to provide more information relative to other conditions, it was also possible that the specific condition might experience the greatest fatigue effects, thereby actually having shorter substantive narratives. The finding that none of the conditions differed, even though children in the specific and control conditions spoke more in practice, is an important one for investigators because it suggests that incident-specific practice is not too long or cognitively demanding that it shortens the rest of the interview.

Substantive Phase: Disclosure of Multiple Instances. Another finding of great practical importance to professionals who regularly interview children who have repeated experiences of abuse is that younger children in the incident-specific practice condition mentioned multiple incidents earlier in the interview than younger children in other

conditions, and children of both ages in the specific condition required marginally fewer prompts to disclose, according to the predictions of hypothesis 2. The results also showed that roughly two-thirds of the older children in all conditions disclosed immediately. It is the younger age group, however, whose testimonies of abuse are the most fragile, and who need more strategies to improve their narratives. These findings are consistent with the assumptions held by researchers who attempt to transfer source-monitoring skills acquired in training to reports of a target event (Poole & Lindsay, 2001; Thierry & Spence, 2002). Children in the incident-specific practice condition were likely able to recognize the commonalities between their autobiographical repeated event and the Laurier Activities, thus realizing that they should talk about both events in the same way.

The NICHD protocol suggests not asking children about multiple incidents until after the first narrative is completed (Orbach et al., 2000, p. 739). If a child is not asked about multiple incidents, however, and begins the narrative with a script-like description of the abuse, the child is rehearsing the script, or strengthening the gist trace (Brainerd & Reyna, 2004). This process facilitates the integration of script-consistent details into one single script, and speeds up the decay of verbatim information. Younger children in the general condition disclosed later in the interview than would be expected, and more younger children than expected in the control condition had to be asked about multiple incidents. These findings suggest that practice in using scripts to describe repeated events encourages continued use of scripts in the substantive phase, and that describing a single-experience event lowers awareness that the interviewer needs to know that the substantive event was a repeated one.

Substantive Phase: Language Use. Children in the specific and control practice conditions continued to use a greater ratio of episodic to scripted language in their substantive narratives than did children in the general condition, as predicted in hypothesis 3. These findings suggest that there are benefits to a subtle manipulation in which children simply practice the desired task, as in source-monitoring training studies (Poole & Lindsay, 2001; Roberts & Cameron, 2006; Thierry & Spence, 2002), rather than being explicitly told what to do. In a field investigation, interviewers sensitive to language would likely notice the script-like dialogue of children in the general practice condition, thereby sensing earlier that the child might be describing a repeated event. Because there were no differences in the language used by children in the specific and control condition, however, the previously reported finding is especially concerning; children in the control condition used as much episodic language as children in the specific condition, which might lead an interviewer to think that they are talking about one time, and many of the control children had to be asked if the Laurier Activities happened more than once. While it is possible that children in the control condition were only describing one occurrence, their source-accuracy score was not higher than other conditions. The alternate explanation is that the children provided a specific, but amalgamated account of the Laurier Activities, which would strengthen the memory trace for that account, and details provided in that account could be used in later interviews with the child.

As predicted by hypothesis 4, age differences were found in the number of temporal, generic and discriminatory references children provided in the substantive portion of the interview, which is consistent with Bauer's (2006) review of

developmental changes in children's event memory. Older children in the current study (7- to 8-years old) are beginning to learn story form, and what elements are important to delivering a relevant narrative. Contrary to expectation, no condition differences or interactions were observed. As can be seen from the means in Table 3, however, the 5- to 6-year olds in the specific condition provided more of all of these references than did the 5- to 6-year olds in the other two conditions. The overall amount of these references was quite low across all children.

Substantive Phase: Recall. Consistent with hypothesis 5, that older children would report overall more information than younger children, the 7- to 8-year olds reported more items and more instantiations than did the 5- to 6-year olds. Similar findings have been observed both in field and analogue studies (e.g., Farrar & Goodman, 1992; Lamb et al., 2003). Younger children in the specific condition, however, reported both more items and more instantiations than younger children in the other conditions, even though the total number of words in their narratives did not differ by condition. Thus, incident-specific practice did in fact encourage the younger children in our sample to report more "forensically relevant" information. Investigators require techniques that are easy to use, and non-suggestive, that enhance the amount of information that young children report. Children's responses to open-ended questions tend to be quite accurate but investigators have the perception that these questions do not encourage enough forensically relevant details in comparison to specific questions (Guadagno, Powell & Wright, 2006), leading them to use more suggestive methods which can be damaging to children's reports (Sternberg et al., 2001).

Contrary to hypothesis 6, that children would report more fixed than variable details (both items and instantiations), children in the current study recalled significantly more variable items and instantiations. Several possibilities for explaining this finding exist. Of the 17 items in the Laurier Activities, five were fixed and seven were variable. Therefore, there is a greater chance of reporting a variable item. The seven variable items naturally had four different instantiations, thereby creating 28 possible instantiations to report that could be classified as variable. The hypothesis that children would provide more fixed than variable details was based upon the assumption that children would only report one instantiation of a variable detail per narrative (e.g., saying that there was a leaf badge). Children often reported additional instantiations of the same item, however (e.g., saying, in list-like fashion, that there were leaf badges, jellybean badges, and feather badges). Only when children very clearly attributed extra instantiations to occurrences other than the one being described, were the instantiations *not* counted as part of their narrative.

Substantive Phase: Difference References. It was expected that practice in describing specific instances of repeated events would encourage children to consider 'what is different' across instances of repeated events, and that this skill would carry over to their substantive narratives. Consistent with hypothesis 7, older children, and children in the specific condition made more references to differences across sessions of the Laurier Activities. More specifically, however, it was once again the younger children in the specific condition that were outperforming the rest of the 5- to 6-year olds. Similar results were found for children's accuracy regarding these difference references. This finding is important in light of Roberts and Powell's work (2003) in which 5- to 6- and 7-

to 8-year olds took part in a very similar repeated event as the one used in the current study ("Deakin Activities"), although without the manipulation of practice phase. Close to one-third (27%) of 5- to 6-year olds were unable to say what was different about the 'target event' to any of the other times.

Substantive Phase: Labels. Moving more specifically now to children's reporting of one instance of the Laurier Activities, hypothesis 8 predicted that older children would provide more of their own labels spontaneously, while younger children would require more interviewer-generated labels. Age differences were not found when the sample was collapsed across condition. When examining the conditions separately, however, it was observed that older children in the control condition required more interviewer labels than would be expected, whereas the younger children in that condition generated more of their own labels than would be expected. Although this result is somewhat unusual for the younger children, it should be noted that child-provided labels were equivalent across control and general conditions ($n = 8$), and slightly higher for the specific condition ($n = 14$). Overall, 73% of all labels were generated by the interviewer.

Older children did have more unique labels, and younger children more non-unique labels than would be expected. Hypothesis 9 predicted that children in the incident-specific condition would have more unique labels than other groups, because they were expected to recall more differences across occurrences, which they did. Only the 5- to 6-year old children in the specific condition had slightly more unique labels than other 5- to 6-year olds, but no condition differences were found for the older children. Once again, a small benefit of incident-specific practice for younger children was observed.

Overall success in both generating a label, and having that label be unique, was not among the *a priori* hypotheses of the current study. When assessing source accuracy, however, it was noted that there were fewer 5- to 6-year olds in the general condition with source scores than children in other conditions. Analyses demonstrated that less than half of the younger children in the general condition were successful at achieving a label for their narrative that could be used to determine accuracy. This finding suggests that general practice, which does appear to encourage a focus on similarities, could be detrimental to young children in forensic interviews who are required to uniquely identify one instance of a repeated event.

Substantive Phase: Source Accuracy. It was expected that older children would be only slightly more accurate than younger children in attributing variable and Hi/Lo details to the correct occurrence, because while younger children should have more difficulty making source-monitoring decisions (Johnson et al., 1993), older children should have stronger gist traces, and a larger pool of gist-consistent instantiations from which to choose the correct one (Brainerd & Reyna, 2004). Analyses demonstrated that older children were in fact only marginally more accurate than were younger children. No effect of condition or interaction was observed. Thus, even though younger children in the specific condition provided more items and more instantiations than did other younger children, their accuracy scores did not differ. Other researchers who study techniques to enhance the amount of information that young children can provide have expressed the concern that acceptable levels of accuracy must be maintained in balance with increased amounts of information (e.g., Elischberger & Roebers, 2001).

Consistent with hypothesis 10, all children more accurately attributed the source of Hi/Lo instantiations in comparison to variable instantiations. Variable instantiations change at each session of the Laurier Activities and thus are very difficult to attribute to the correct session. Hi/Lo instantiations remain the same for three occurrences, and are different on one occurrence, making both instantiations very memorable.

Contrary to hypothesis 11, younger children did not make more external intrusions. Overall, the rate of external intrusions was very low, demonstrating that even the young children had likely built up a script of the Laurier Activities, and solidifying the finding that children with repeated experiences, although they can be confused about what happened on which occasion, rarely report things that never happened when questioned non-suggestively (Roberts & Powell, 2001).

Although differences among the practice conditions, and between the age groups were modest in the current study, taken together they point to the benefits of incident-specific practice after a delay, especially for the young children. The finding that a technique enhances recall and overall performance for younger, but not older, children is a common theme in developmental literature when older children have already acquired, and are able to use, the strategy or skill being trained (Bjorklund, Miller, Coyle, & Slawinski, 1997). Research on the NICHD protocol interview itself, on which the current study's interviewing was modelled, has shown greater benefits of the technique for younger children (4- to 8-year olds) than for older children (9- to 12-year olds) (Sternberg et al., 2001).

One reason that large differences between the two practice conditions may not be statistically significant is that verbatim traces may have decayed to the point where they

cannot be accessed at the time of the interview. Fuzzy trace theory suggests that verbatim traces can decay quite quickly, even over a matter of a few days, and that they decay faster in younger children (Brainerd & Reyna, 2004), which suggests that the benefits observed for the 5- to 6-year olds in the incident-specific condition may be more substantial than can be determined statistically. It is also possible, according to the schema-confirmation-deployment model (Farrar & Goodman, 1992) that the younger children took longer to build up their script of the Laurier Activities, such that they required all four sessions to simply confirm their script, thus assimilating many deviations, and may not have fully reached the deployment stage at the time of the interview. Finally, it is also possible that children of both ages had intact verbatim traces, and they had developed functioning scripts, but that their source-monitoring skills were poor and therefore could not decide which detail occurred during which session, as would be suggested by the source-monitoring framework (Johnson et al., 1993). Research has shown that children *and* adults have difficulty with source-monitoring when events are highly similar (e.g., Lindsay et al., 1991), such as with instances of repeated events.

All findings taken together, benefits were observed for children in the incident-specific practice condition. Additionally, even though their narratives were neither more nor less accurate, they disclosed multiple incidents earlier, used more specific language, and recalled more differences across sessions, making their narratives appear more concise and confident. It is likely that these narratives would appear more credible to blind observers than the narratives of children in the other conditions, because research has shown that children who testify in a confident manner receive higher credibility judgments from mock jurors (Schmidt & Brigham, 1996). Future research is planned to

examine whether undergraduate students will rate the narratives of children in the incident-specific practice condition as being more credible.

Limitations

Several limitations of the current study should be addressed. Although there did not appear to be effects of gender or number of interviewer prompts in practice on any of the main analyses, there were differences among conditions on these variables. Although practice in describing *two* instances of a repeated event was integral to the incident-specific practice condition, the number of prompts would have been equal across conditions without the second-incident practice. Future research is needed to further clarify our suggestion that since covarying the number of prompts in practice did not change major findings, it is likely to be the type, and not the amount, of practice that had effects.

The problems with unequal variances observed in many of the analyses reported here is likely to be a common, although often unreported, problem in studies that examine developmental differences among children. Younger children typically evidence greater variability than do older children in cognitive and linguistic tasks (e.g., Kopp, 1978), and this problem is unlikely to be solved by increasing sample size. The number of older children would have to increase by the same amount, and thus the variability of both groups would decrease, but not become more homogenous. As with all statistical analyses in which there are problems with the data, results should be interpreted with caution.

Finally, in an ideal world, unique labels would have been generated for all children in the current study, making sample sizes equal across conditions and ages for

analyses of accuracy. Doing so, however, would have introduced a level of artificiality into the interview process, because labels would be created by the interviewer based on information that had not been provided by the child. Differences between ages and among conditions would be less informative, because the children who would have been the least successful at achieving a unique label on their own, would essentially be engaging in cued recall for the session probed by the label. As discussed by Roberts and Powell (2003a), much more research is needed concerning how to help children generate unique labels, such that investigators can feel confident that both they and the child are discussing the same occurrence of a repeated event.

Implications and Future Directions

Certainly more research on types of practice would be necessary before implementation of the technique. Firstly, since interviewers do not always know at the outset of an investigation whether a child has been abused on multiple occasions, it would be important to determine how practice in describing an autobiographical repeated event might affect the narratives of children who have experienced only a single instance of abuse.

Additionally, because benefits of incident-specific practice were primarily observed for the younger children in the sample, it would be highly relevant to test the technique with preschool aged children whose scripts would be more "skeletal" (Hudson et al., 1992), less likely to be fully confirmed (Farrar & Goodman, 1992), and who would have weaker verbatim traces (Brainerd & Reyna, 1998), and less ability to make source-decisions (Drummey & Newcombe, 2002) than even the 5- to 6-year olds in the current study.

A careful comparison of children's descriptions of different types of repeated events would also be relevant. Perhaps practicing discussing events such as holidays, which occur with regular frequency but are spaced relatively far in time, and have a very strong collective knowledge base (e.g., most people could describe "Christmas" even if they do not celebrate it), is practically different from describing a more personal, more frequently occurring event (e.g., swimming lessons).

Another issue relevant to future research is to attempt to systematically test the theories and frameworks reviewed. The current study was not designed to determine whether any one theory was more adequate than another in understanding how incident-specific practice might affect the behaviour of children with repeated-event experience. Rather, the current research was born from a desire to study an under-examined area of the NICHD interview protocol, and hypotheses were informed by script theories, fuzzy-trace theory and the source-monitoring framework. It was earlier mentioned that results showing only increases in episodic language use but not accuracy, as in the current study, would lend more support to script theories, while increases in episodic language use *and* subsequent increases in accuracy would likely be predicted by fuzzy-trace theory and the source-monitoring framework. Because the study was not designed to explicitly test these theories against each other, however, it cannot be known for example, whether intact verbatim traces persisted at the time of test, or whether children were able to make effective source-decisions. Future research could attempt to 1) ensure that all children have fully encoded each event (maximizing the likelihood that all children could potentially recall an intact verbatim trace) and 2) use source-monitoring training techniques, to enhance the probability that all children can make source-decisions, to

some degree. If the effects of encoding, and decisions made at retrieval, could be equalized across all children, it could be determined whether incident-specific practice alone has any further benefits. In addition, more confident conclusions could be made as to the utility of the theories reviewed in this research.

One interesting methodological implication arose in the current study. Due to issues with availability, six interviewers were required on the research team to interview all children within the required delay. All interviewers did run children in all practice conditions, and children in each age group were balanced across interviewers as much as it was possible. All interviewers received the same type of training, and all were given as much training as was needed until the author was satisfied with performance. All interviewers received additional re-training throughout the course of the study, both orally, and with written assessments of their transcribed interviews. Yet, subtle interviewer differences in style were apparent. While these differences could be considered a limitation of the current study, they also enhance the generalizability of the findings. New research in the field of child interviewing has been focused on the bi-directional influences of the child-interviewer interaction (e.g., Gilstrap & Papierno, 2004). Every child-interviewer dyad can have a unique social dynamic. For example, Gilstrap and Papierno (2004) found that the relationships observed between children's social competence and their levels of verbosity, cooperation and acquiescence were mediated by interviewer behaviour (most specifically, use of leading questions). What is especially interesting is that these effects were observed in a sample of 38 children, interviewed by 38 trained police investigators.

Conclusion

In conclusion, incident-specific practice appears to benefit younger children more so than older children, although both age groups experienced some benefits over other conditions. This type of practice encourages behaviours and recall that would be very relevant to field investigators working with young children who have multiple event experiences, such as earlier disclosure of multiple incidents, greater recall and more episodic narratives that are not less accurate, and greater recognition of differences across highly similar repeated events. Future studies should examine the effects of incident-specific practice for children with single experiences, as well as for children with repeated experiences over longer delays (between incidents, and between the final incident and the interview). Implementation of this procedure (versus standard practice currently in use) could be used in the field and comparisons could be made on most variables (except accuracy) between actual forensic interviews and this analogue study. Additionally, future research should examine whether children in the incident-specific condition provided narratives that appear more credible to participants than children in the other conditions.

Table 1

Observed and Expected Results for Location of Disclosure, by Age and Condition

			Location of Multiple Incidents Disclosure				
			Labelling	Immediate	Breadth/ Depth	Other	Total
Age 5-6							
Condition	Control	Observed	7	2	9	0	18
		Expected	3.1	3.1	10.6	1.2	18
	General	Observed	3	3	11	3	20
		Expected	3.4	3.4	11.7	1.4	20
	Specific	Observed	0	5	14	1	20
		Expected	3.4	3.4	11.7	1.4	20
Total			10	10	34	4	58
Age 7-8							
Condition	Control	Observed	2	15	3	0	20
		Expected	1.7	14	4.3	0	20
	General	Observed	2	11	7	0	20
		Expected	1.7	14	4.3	0	20
	Specific	Observed	1	16	3	0	20
		Expected	1.7	14	4.3	0	20
Total			5	42	13	0	60

Table 2

Ratio of Episodic to Scripted language in Substantive Phase

	Condition		
	Specific	General	Control
Age			
5-6	.83 (.17)	.59 (.30)	.90 (.13)
7-8	.92 (.08)	.74 (.26)	.86 (.18)

Note: Standard deviations are in parentheses.

Table 3

Mean raw totals and standard deviations by Age and Condition for number of Temporal, Generic, and Discriminatory references provided in Substantive Phase

		Condition		
		Specific	General	Control
	Age			
Temporal	5-6	1.80 (2.46)	.70 (1.17)	1.06 (1.39)
References	7-8	3.90 (3.08)	2.40 (2.78)	3.10 (4.84)
Generic References	5-6	3.30 (3.08)	2.85 (3.47)	1.28 (1.87)
	7-8	4.25 (4.18)	4.25 (3.23)	4.70 (4.70)
Discriminatory	5-6	3.20 (2.95)	1.90 (2.40)	1.28 (1.87)
References	7-8	3.55 (2.78)	3.95 (3.25)	3.10 (2.73)

Note: Standard deviations are in parentheses.

Table 4

Mean raw totals and standard deviations by Age and Condition for Accurate Difference References to Variable and Hi/Lo instantiations

		Condition		
		Specific	General	Control
	Age			
Variable	5-6	1.80 (1.36)	.65 (1.04)	.67 (.77)
	7-8	1.90 (1.29)	1.60 (1.43)	1.60 (1.05)
Hi/Lo	5-6	.75 (.79)	.15 (.37)	.22 (.55)
	7-8	.45 (.85)	.85 (.67)	.70 (.73)

Note: Standard deviations are in parentheses.

Appendix A

Interview Opening:

"It's my job to talk to children to see what they remember about things. I heard that someone visited your class and did the Laurier Activities with you, so I'll be asking you some questions about that. I usually write things down to help me remember what I need to ask you about."

"But before we talk about the Laurier Activities, I want to get to know you a little better. Your (parent/guardian) was telling me that you _____(parent-provided event)."

Incident-specific practice

"Some things we do when we _____are always the same, and some things are different. I'd like you to tell me about *the time you remember best when* you _____. Tell me everything you can remember about the *time you remember best when* you _____from the very beginning to the very end. I don't want you to leave anything out. And remember I only want you to tell me about the *time you remember best when you* _____."

Prompt for a second occurrence:

"Now I would like you to tell me all about another time when you _____, every little detail from the very beginning to the very end."

General practice

"I'd like you to tell me *all about* when you _____. Tell me everything you can remember _____ from the very beginning to the very end. I don't want you to leave anything out. And remember I want you to tell me *all about* when you _____."

Prompt for more information

"Tell me *more* about what happens when you _____."

Control practice

"But before we talk about the Laurier Activities, I want to get to know you a little better. Your (parent/guardian) was telling me that you _____(parent-provided event) for the first time.

I'd like you to tell me about when you _____ for the first time. Tell me everything you can remember _____ from the very beginning to the very end. I don't want you to leave anything out. And remember I want you to tell me everything that happened when you _____."

Prompt for more information

"Tell me *more* about what happened when you _____."

All children get at least one **"What else can you tell me about _____" prompt**, and an unlimited array of non-verbal/minimal encouragers.

Tell me more about _____ [something that child has mentioned].

Tell me everything about _____ [something that child has mentioned].

OK. It sounds like you had a really great time _____.

Appendix B

The following sample protocol is not an exhaustive list of instructions and prompts for interviewers, but is intended to give the reader a general sense of what types of questions will be asked and how they were worded.

The interview is divided into 5 sections;

PT1: Practice (5-10 minutes approx) where the child recalls parent-provided event (Appendix A)

PT2: Narrative about one of the occurrences (15 minutes maximum)

PT3: Narrative about the second occurrence (15 minutes maximum)

PT4: Difference questions (10 minutes approx)

PT5: Closure.

These times are recommended only, except that the times for the narratives cannot exceed the maximum time-frames.

PT 2 (1st narrative):

Note that narrative sections are unstandardized. They are to be non-leading and non-suggestive. The interviewer is allowed to draw on the pool of techniques suggested by the NICHD protocol (e.g., only asking for further information about things the child has already mentioned).

"Now it's time to talk about the Laurier Activities. Remember, I wasn't there when you did the Laurier Activities and it's really important that I know what happened. I need to know every little detail, so tell me everything you remember about the Laurier Activities from the very beginning to the very end."

Sample Probes:

- Tell me more about _____ (something mentioned by child)

  What else can you tell me about _____?

  - *Probing breadth*

- "What happened then", "What else can you tell me about the Laurier Activities?", "What else can you tell me that happened when the lady did the Laurier Activities with you?"

  - *Probing depth*

- "Tell me more about the part when _____ [something that child has mentioned]."
- "Tell me everything about the part when_____."
- "Tell me every detail about the part when _____ from the very beginning to the very end."
- "What else can you tell me about____?"

Probing Multiple Events

If child makes clear by the use of language that there were multiple events (i.e., One day we got X, one day we got Y) say **"It sounds like you did the Laurier Activities more than one time. Tell me about the time you remember best."**

If child suggests by the use of language that there were multiple events but that it is not clear (i.e., the child talks in the present tense, or says sometimes get X, sometimes Y), say **"I'm a little confused. Did you do the Laurier Activities one time or more than one time?".**

Providing Encouragement

Note that this narrative section can be tiring for the child so give the child some reinforcement for the process where needed: "You've told me a lot about the Laurier Activities and I understand it much better now", "You're being very helpful", "I can see that you've been thinking hard".

PT 3 (2nd narrative):

"Tell me about another time that you did the Laurier Activities."

[If child provides minimal information, says "I don't know", or says it was "the same", ask **"It's really important that I know all about the times that you did the Laurier Activities. Think carefully and tell me about another time that you did the Laurier Activities. It's really important that you tell me every single detail about another time you did the Laurier Activities".**

PT 4 (difference questions)*:

This questioning is asked for all those children who have reported about two occurrences of the event.

"So now you've told me about the____ time and the ___ time. I just need to look at my notes to make sure that I haven't forgotten to ask you any questions."

'You mentioned there was a ____ [general item, as non-suggestively as possible; try and stick to child's description] in the ___ time as well as the ____ time. Was there anything different about the _____ each time? Tell me what was different about the _____?'(Interviewer tries to elicit which instantiation is associated with each label).

* Note: Analyses not included in current paper

PT 5 (closure):

"You've really told me an awful lot about what happened when you did the Laurier Activities. Thank you for talking to me. You were very helpful."

Appendix C

FULL SET OF TARGET ITEMS

	ITEM	A	B	C	D
1.	Children sit on X	cardboard	blue mat	garbage bag	newspaper
2.	Cloak of leader	Red	Yellow	green	blue
3.	Fox's name	Boo	Kip	Pop	Jo
4.	Noisy animal	Seal	Penguin	walrus	polar bear
5.	Warm-up activity	Dance	wiggle fingers	touch toes	jump
6.	Source of story	Cupboard	leader wrote	Bought from Internet	library
7.	Content of story	Dog in city	Winter	party	Boat
8.	Bookmark	Pink hearts	Black triangles	Purple squares	Orange circles
9.	Utensil to note story	Pencil	Crayon	chalk	Felt pen
10.	Puzzle	Juggling	Bike	tightrope	Car
11.	Music/scene for resting	Beach	Kites/park	birds	Rain
12.	Part of body is relaxed	Legs	Nose	stomach	Arms
13.	Method of getting refreshed	baby wipe	cool drink (water)	Hand sanitizer	fan
14.	Theme of sticker	Rocket	Dinosaur	apple	Ball
15.	Container with stickers	Box	Purse	envelope	Jar
16.	Next stop	to movie	walking a dog	visiting friend in hospital	birthday party
17.	Badge	Jelly bean	Fluffy Pink	Leaves and Bark	Buttons

Counterbalancing Procedure

Item Sets and Order of Administration:

In all Age x Practice conditions, half of the children are in Group 1 and the remainder in Group 2

Group 1 children receive the Occurrences CDBA in that order. Group 2 children receive the Occurrences ABDC in that order.

Status of Target Items:

There are 17 target items as shown above.

5 of these 17 are *fixed* across the occurrences.

7 of these 17 are *variable* across the occurrences.

5 of these 17 are *Hi/Lo* across the occurrences, meaning that for three of the occurrences the item is fixed, and for one it varies. The occurrence which contains the Lo frequency item is counterbalanced by rotating occurrences 1 through 4 for Group 1, and rotating occurrences 4 through 1 for Group 2.

The actual items that are assigned to fixed, variable, and Hi/Lo status differ for each group and are shown below. The occurrence which constitutes the low frequency instantiation for the Hi/Lo items is shown in parentheses:

	Group 1	**Group 2**
Fixed	1, 4, 6, 12, 16	7, 9, 10, 11, 14
Variable	2, 5, 7, 9, 14, 15, 17	3, 4, 6, 8, 13, 16, 17
Hi/Lo	3 (1), 8 (2), 10 (3), 11 (4), 13 (1)	1 (4), 2 (3), 5 (2), 12 (1), 15 (4)

Item 17 (badge) is counterbalanced for item sets and order of administration, but is not counterbalanced for status of item. It is administered as a variable item for both groups. This is because it is the only item that is available to the children throughout the entire event and as such it should serve as a good tag for the events. We therefore left it variable to see whether the children would use this tag.

References

Abelson, R.P. (1981). Psychological status of the script concept. *American Psychologist,* 36, 715-729.

Bauer, P.J. (2006). Event memory. In D. Kuhn, R.S. Siegler, W. Damon, & R.M. Lerner, (Eds.), *Handbook of child psychology: Vol.4.Cognition, perception and language* (6th ed., pp. 373-425). Hoboken, NJ: Wiley.

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

Brainerd, C.J., & Gordon, L.L. (1994). Development of verbatim and gist memory for numbers. *Developmental Psychology,* 30, 163-177.

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

Brainerd, C.J., Reyna, V.F., Howe, M.L., & Kingma, J. (1990). The development of forgetting and reminiscence. *Monographs of the Society for Research in Child Development,* 55, 1-111.

Brainerd, C.J., & Reyna, V.F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review,* 100, 42-67.

Brainerd, C.J., & Reyna, V.F. (1998). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology,* 71, 81-129.

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

Connolly, D. A., & Lindsay, D.S. (2001). The influence of suggestions on children's reports of a unique experience versus reporting of a repeated experience. *Applied Cognitive Psychology,* 15, 205-223.

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

Elischberger, H.B., Pipe, M.-E., Orbach, Y., & Lamb, M.E. (2005, May). *Do young children rely on scripts in recounting multiple instances of abuse?* Paper presented at the 17th Annual Convention of the American Psychological Society, Los Angeles, CA.

Elischberger, H.B, & Roebers, C.M. (2001). Improving young children's free narratives about an observed event: The effects of non-specific verbal prompts. *International Journal of Behavioral Development,* 25, 160-166.

Erskine, A., Markham, R., & Howie, P. (2002). Children's script-based inferences: Implications for eyewitness testimony. *Cognitive Development,* 16, 871-887.

Evans, A. & Roberts, K. (2007). *Can paraphrasing improve the quality of reports from young child witnesses?* Manuscript submitted for publication.

Farrar, M.J., & Boyer-Pennington, M.E. (1999). Remembering specific episodes of a scripted event. *Journal of Experimental Child Psychology,* 73, 266-288.

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

Fivush, R., Hudson, J., & Nelson, K. (1984). Children's long-term memory for a novel event: An exploratory study. *Merrill-Palmer Quarterly,* 30, 303-316.

Gilstrap, L.L., & Papierno, P.B. (2004). Is the cart pushing the horse? The effects of child characteristics on children's and adults' interview behaviours. *Applied Cognitive Psychology,* 18, 1059-1078.

Graesser, A.C., Gordon, S.E., & Sawyer, J.D. (1979). Recognition memory for typical and atypical actions in scripted activities: Tests of a script pointer + tag hypothesis. *Journal of Verbal Learning and Verbal Behavior,* 18, 319-332.

Guadagno, B.L. & Powell, M.B. (2006). *A qualitative examination of police officers' questioning of children about a repeated event.* Manuscript submitted for publication.

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

Hudson, J.A. (1990). Constructive processing in children's event memory. *Developmental Psychology*, 26, 180-187.

Hudson, J.A., Fivush, R., & Kuebli, J. (1992). Scripts and episodes: The development of event memory. *Applied Cognitive Psychology. Special Issue: Memory in everyday settings,* 6, 483-505.

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.

Jones, D.C., Swift, D. J., & Johnson, M.A. (1988). Nondeliberate memory for a novel event among preschoolers. *Developmental Psychology,* 24, 641-645.

Kopp, C.B. (1978). Individual differences and interventions for infants and young children. *Journal of Pediatric Psychology,* 3, 145-149.

Lamb, M.E., Sternberg, K.J., Esplin, P.W., Hershkowitz, I., Orbach, Y., & Hovav, M. (1997). Criterion-based content analysis: A field validation study. *Child Abuse & Neglect,* 3, 255-264.

Lamb, M.E., Sternberg, K.J., Orbach, Y., Esplin, P.W., Stewart, H., & Mitchell, S. (2003). Age differences in young children's responses to open-ended

invitations in the course of forensic interviews. *Journal of Counselling & Clinical Psychology,* 71, 926-934.

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

Mandler, J.M. (1979). Categorical and schematic organization in memory. In R. Puff (Ed.) *Memory Organization and Structure* (pp. 259-299). New York: Academic Press.

McCauley, M., Langrock, A.M., & Redmond, S. (2007, April). "Before we begin, I'd like to get to know you better": The importance of establishing good rapport. In Langrock, A.M., & McCauley, M. (Chairs), *The value of good rapport techniques in forensic and clinical interviews.* Symposium conducted at the biennial meeting of the Society for Research in Child Development, Boston, MA.

McNichol, S., Shute, R., & Tucker, A. (1999). Children's eyewitness memory for a repeated event. *Child Abuse & Neglect*, 23, 1127-1139.

Moravcsik, J.E. & Healy, A.F. (1995). Effect of meaning on letter detection. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 21, 82-95.

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

Orbach, Y., Hershkowitz, I., Lamb, M.E., Sternberg, K.J., Esplin, P.W., Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse & Neglect,* 24, 733-752.

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

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

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

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.

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

Reyna, V.F., Holliday, R., & Marche, T. (2002). Explaining the development of false memories. *Developmental Review,* 22, 436-489.

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

Roberts, K.P., & Cameron, S. (2006). *"I remember, I just can't remember where": Training children to discriminate between memories of highly similar events.* Manuscript submitted for publication.

Roberts, K.P., Lamb, M.E., & Sternberg, K.J. (2004). The effects of rapport-building style on children's reports of a staged event. *Applied Cognitive Psychology,* 18, 189-202.

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

Roberts, K. P., & Powell, M. B. (2003a). *Children's ability to effectively label different instances of a repeated event.* Paper presented at the biennial meeting of the Society for Research in Child Development, Tampa, FL.

Roberts, K. P., & Powell, M. B. (2003b). *Children's awareness of the distinctiveness of different sources of information.* Paper presented at the biennial meeting of the Society for Research in Child Development, Tampa, FL.

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

Roberts, K.P., & Powell, M.B. (2006b). [How effective are recommended interviewing guidelines in eliciting narrative accounts of specific occurrences of a repeated event?]. Unpublished manuscript.

Roberts, K. P. & Powell, M.B. (in press). The roles of prior experience and the timing of misinformation presentation on young childrens' event memories. *Child Development.*

Schmidt, C.W., & Brigham, J.C. (1996). Jurors' perceptions of child victim-witnesses in a simulated sexual abuse trial. *Law and Human Behavior,* 20, 581-606.

Siegal, M. (1979). Children's comprehension of verbally stated rules: Knowing that the meaning of sometimes is different from always or never. *Psychological Report,* 44, 932-934.

Sternberg, K.J., Lamb, M.E., Orbach, Y., Esplin, P.W., & Mitchell, S. (2001). Use of a structured investigative protocol enhances young children's responses to

free-recall prompts in the course of forensic interviews. *Journal of Applied Psychology,* 86, 997-1005.

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