# THE INFLUENCE OF DIRECTED FORGETTING ON AUTOBIOGRAPHICAL MEMORY: THE ROLE OF DEVELOPMENT, INDIVIDUAL DIFFERENCES, AND REPEATED INTERVIEWS

by

Heidi M. Gordon
B.A. (Honours), Queen's University, 2003
M.A., Simon Fraser University, 2005

DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

In the Department of Psychology



SIMON FRASER UNIVERSITY

Spring 2009

# APPROVAL

| | |
|---|---|
| **Name:** | Heidi M. Gordon |
| **Degree:** | Doctor of Philosophy (Department of Psychology) |
| **Title of Thesis:** | The Influence of Directed Forgetting on Autobiographical Memory: The Role of Development, Individual Differences, and Repeated Interviews |
| **Chair:** | Dr. Mark Blair<br>Assistant Professor |
| **Examining Committee:** | Dr. Deborah Connolly<br>Senior Supervisor<br>Associate Professor |
| | Dr. Cathy McFarland<br>Supervisor<br>Professor |
| | Dr. Thomas Spalek<br>Supervisor<br>Professor |
| **Internal Examiner:** | Dr. Wanda Cassidy<br>Associate Professor<br>Faculty of Education |
| **External Examiner:** | Dr. Elyse Lehman<br>Professor Emerita<br>Department of Psychology<br>George Mason University |
| **Date Approved :** | April 2, 2009 |

SFU SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

# STATEMENT OF ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Bennett Library
Simon Fraser University
Burnaby, BC, Canada

# ABSTRACT

Children often participate in the criminal justice system as victims of childhood sexual abuse (CSA) and, in most cases, these children do not disclose the abuse for some time. During this period of non-disclosure, CSA victims may avoid thinking about or discussing their abuse experiences. Some scholars argue that this may lead to a directed forgetting (DF) effect, whereby later recall of the unrehearsed memories becomes more difficult, but limited empirical evidence exists to support this position. The present research employed an innovative application of the directed forgetting (DF) protocol to memory for a personally experienced event. Two hundred and forty-six participants from four age groups (grade one, grade three, grade five, and university) completed an activity session during which the experimenter issued list method DF instructions. Interviews consisting of executive functioning (EF), recall, and recognition tasks occurred six and seven days later. Findings demonstrated typical costs of the forget cue: Participants recalled fewer details from the first two activities when they were forget-cued than when they were remember-cued. However, follow-up tests revealed that only children in grade five and adults demonstrated these DF costs. Performance on EF tasks did not predict DF effects. In addition, those participants who recalled the activity session in the first interview did not demonstrate reduced DF costs during the second interview. This study marks the beginning of an important area of research; implications for the application of DF to memory for CSA are discussed in light of the present findings.

# DEDICATION

For my family, the love and support of which reaches me from any distance.

# ACKNOWLEDGEMENTS

I would first like to extend my deepest gratitude to my supervisor, Dr. Deborah Connolly. Her tireless support and unwavering belief in my abilities has sustained me these last six years. Thanks to her mentorship, I am completing graduate school with the skills, maturity and confidence required for the next stage of my career. In addition, I would like to thank my committee members, Dr. Tom Spalek and Dr. Cathy McFarland, for their insight and wisdom throughout this process. My research was supported by a doctoral fellowship from the Social Sciences and Humanities Research Council (SSHRC) and a research grant from the Human Early Learning Partnership (HELP).

Thanks to Jen Lucyk, Sukhraj Sihota, Aline Levasseur, Iris Paradela, and Shannon Plumpton for leading activity sessions and conducting interviews. I would also like to acknowledge the many principals, teachers, parents, children, and undergraduate students for their participation in this project. The combined efforts of these individuals facilitated timely completion of my data collection, coding, and analysis.

Special thanks to Derek, who inspires me; and to the Burnaby girls, who taught me the true meaning and value of friendship.

# TABLE OF CONTENTS

**Approval** .......... ii

**Abstract** .......... iii

**Dedication** .......... iv

**Acknowledgements** .......... v

**Table of Contents** .......... vi

**List of Figures** .......... viii

**List of Tables** .......... ix

**Introduction** .......... 1

*Directed Forgetting* .......... 2
*Directed Forgetting and Children* .......... 11
*Directed Forgetting and Autobiographical Memory* .......... 15
*Executive Functioning* .......... 18
*Repeated Interviews* .......... 21
*The Present Research* .......... 25

**Method** .......... 27

*Participants* .......... 27
*Materials* .......... 27
Shifting. .......... 27
Updating. .......... 28
Working memory. .......... 29
Inhibitory control. .......... 30
*Design and Procedure* .......... 31
Activity sessions .......... 31
Memory interviews. .......... 32
Coding. .......... 34

**Results** .......... 36

*Directed Forgetting* .......... 36
Interview 1. .......... 36
Interview 2. .......... 37
*Executive Functioning* .......... 41
Manipulation Checks. .......... 41
Developmental Differences .......... 43
Regression Analyses. .......... 44
*Repeated Interviews* .......... 45

Release from inhibition. .......... 45
Report consistency. .......... 46

**Discussion** .......... **48**

*Directed Forgetting* .......... 48
Theoretical explanations. .......... 48
Developmental differences. .......... 53

*Executive Functioning* .......... 56

*Repeated Interviews* .......... 58

*Limitations and Future Directions* .......... 60

**Conclusion** .......... **65**

**Appendices** .......... **84**

*Appendix A. Example of first interview* .......... 85

*Appendix B. Example of second interview* .......... 86

**References** .......... **90**

# LIST OF FIGURES

Figure 1. *Mean (Standard Error Bars) Recall of List 1 Details as a Function of Age and DF Condition for Interview 1* .......... 67

Figure 2. *Mean (Standard Error Bars) Recall of List 1 Details as a Function of Age and DF Condition for Interview 2* .......... 68

# LIST OF TABLES

Table 1. *List of Critical Details* ........ 68

Table 2. *Means (SDs) of accurate List 1 and List 2 recall in Interview 1* ........ 70

Table 3. *Means (SDs) of accurate List 1 and List 2 recall in Interview 2* ........ 71

Table 4. *Means (SDs) of correct "Yes" Responses in Recognition* ........ 73

Table 5. *Descriptive statistics for Trail Making, Keep Track, Digit Backwards, and Stroop task performance* ........ 75

Table 6. *Correlations between Trail Making (TA, TB), Keep Track (KT), Digit Backwards (DB), and Stroop (SW, SC, SCW) tasks for Grade One students* ........ 76

Table 7. *Correlations between Trail Making (TA, TB), Keep Track (KT), Digit Backwards (DB), and Stroop (SW, SC, SCW) tasks for Grade Three students* ........ 77

Table 8. *Correlations between Trail Making (TA, TB), Keep Track (KT), Digit Backwards (DB), and Stroop (SW, SC, SCW) tasks for Grade Five students* ........ 78

Table 9. *Correlations between Trail Making (TA, TB), Keep Track (KT), Digit Backwards (DB), and Stroop (SW, SC, SCW) tasks for university students* ........ 79

Table 10. *Multiple Regression Model* ........ 80

Table 11. *Correlations between Predictors and Difference Scores (DIFF) for FR Participants* ........ 81

Table 12. *Mean Proportions (SDs) Consistency for List 1 Details* ........ 82

Table 13. *Mean Proportions (SDs) Consistency for List 2 Details* ........ 83

# INTRODUCTION

Historically, the legal community deemed children inherently unreliable witnesses and several barriers prohibited children from providing evidence (e.g., Bala, 1999). Substantial legal reform has occurred over the last 30 years, documenting a commitment to the protection of children and increased accessibility of the criminal justice system to child witnesses (e.g., Bill C-2, 2005 in Canada). Consequently, more children now participate in various stages of the investigative and adjudicative processes, and legal and psychological professionals are increasingly aware of child witness needs. Child sexual abuse (CSA) is a frequent cause of children's involvement in the criminal justice system and, due to limited corroborating evidence in these cases, the child's complaint is often the only available prosecutorial evidence (Bala, Lee, & McNamara, 2001). Thus, adjudication is based largely on an evaluation of credibility, and understanding factors that affect the accuracy and completeness of child witness statements is critical (e.g., Bruck, Ceci, & Hembrooke, 2001; Lamb, 1999). When examining the types of errors children make in their memory reports, researchers have focused on children's errors of commission (i.e., fabricating details of an event). However, some have argued that children's memory reports are more likely to contain omissions (i.e., failing to report details of an event) and very little is known about these errors (Goodman, Bottoms, & Rudy, 2001).

The prevalence of omissions in children's reports has important implications for the forensic context to which this work is applied. In a recent review of CSA disclosure patterns, London, Bruck, Ceci, and Shuman (2005) found that approximately two-thirds of children delay disclosing their abuse experiences. There are several characteristics inherent to the abuse experience that may foster delayed disclosure. From the victim's perspective, perceived shame and embarrassment dissuade discussion of or thoughts about the abuse. In addition, perpetrators may encourage non-disclosure by urging the victim to keep the abuse a secret, denying that the abuse occurred, refusing to discuss the abuse, or explicitly telling the victim to "forget that this happened". Thus, to prevent disclosure, victims may exert considerable effort in avoiding rehearsal or retrieval of abuse-related thoughts. Scholars have argued that this may lead to a directed forgetting (DF) effect, whereby later recall of the suppressed memories becomes more difficult (e.g., Anderson & Green, 2001; Epstein & Bottoms, 2002; Koutstaal & Schacter, 1997). However, there is little empirical evidence to support this claim.

One of the primary objectives of this dissertation was to apply the directed forgetting (DF) protocol to memory for a personally experienced event. Specifically, the present research examined three under-explored issues relevant to some child witnesses' experiences: (1) directed forgetting may lead to omissions in autobiographical memory reports, (2) the ability to direct forgetting of an event may develop with age and be related to executive functioning, and (3) repeated interviews may affect the consistency with which children omit details from their autobiographical memory reports. To examine these issues, participants completed a series of activities they were directed either to remember or to forget and later recalled the activities.

## ***Directed Forgetting***

In a typical directed forgetting (DF) experiment, participants are instructed to forget some information presented to them and to remember the remaining information. This is often in the form of words in a list or story. There are two common methods of issuing the forget/remember cues: the item method and the list method. Forget cues (e.g., FFFF) and remember cues (e.g., RRRR) are provided after each word in the item method, whereas the forget/remember cues follow each list of words in the list method. The DF manipulation is within-subjects for item method studies and between-subjects for list method studies. Thus, in list method studies, two conditions result wherein half of the participants are instructed to forget List 1 words and remember List 2 words (i.e., FR condition) and the remaining half of the participants are instructed to remember both List 1 and List 2 (i.e., RR condition). After presentation of the words, participants in both item method and list method studies are asked to recall as many items as they can and answer recognition questions about all items. A "directed forgetting effect" occurs when the instructional cues affect memory, and can be measured by either the cost/benefit analysis or the remember-forget difference. Cost/benefit analyses are typically conducted in list method studies: Compared to participants in the RR condition, participants in the FR condition recall fewer List 1 items (i.e., cost of the forget cue) and more List 2 items (i.e., benefit of the forget cue). Using the remember-forget difference, typically conducted for item method studies, a significant difference between number of remember-cued items recalled and number of forget-cued items recalled is indicative of a directed forgetting effect (see MacLeod, 1998 for a review).

A skeptical, and perhaps justified, reaction to the DF effect is disbelief that participants truly forget the forget-cued items and are not merely responding to demand characteristics. However, researchers examining the role of demand characteristics on DF have failed to find evidence of intentional withholding of forget-cued items. For example, Geiselman, Rabow, Wachtel, and MacKinnon (1985; Exp. 4) did not find that providing an incentive at retrieval facilitated recall of forget-cued items. Participants in their study were assigned to one of three incentive conditions: (1) participants received 9 points for recalling remember-cued items and 1 point for recalling forget-cued items, (2) participants received 9 points for recalling forget-cued items and 1 point for recalling remember-cued items, or (3) participants received 5 points for each item recalled. They found significant DF effects in all incentive conditions, suggesting that participants were not responding to demand characteristics. As an extension of this research, MacLeod (1999) argued that the point system developed by Geiselman and colleagues may not have been meaningful to participants, and paid participants $0.50 for every previously unreported forget-cued item recalled during a second recall test. The monetary incentive did not facilitate recall of previously unreported forget-cued items: Participants did not recall more new forget-cued words than new remember-cued words during this second recall attempt. Taken together, these findings do not support the notion that participants intentionally withhold forget-cued items during recall tasks, and efforts to understand DF have focused on other mechanisms.

Interpretations of the DF effect have undergone marked transformations since the protocol's inception (R. A. Bjork, Laberge, & Legrand, 1968). This is understandable given the initially conflicting findings. Although performance on recall tests was generally consistent, in that participants recalled fewer forget-cued words than remember-cued words (e.g., E. L. Bjork & R. A. Bjork, 1996; Geiselman, R. A. Bjork, & Fishman, 1983), recognition performance varied. Some researchers found significant DF effects for recognition tasks (e.g., R. A. Bjork & Geiselman, 1978; MacLeod, 1975) and others did not (e.g., Block, 1971; Elmes, Adams, & Roediger, 1970). Upon further examination, it appeared as though recognition task performance depended on method of cue presentation. Specifically, participants exposed to the item method provided fewer correct responses to recognition questions about forget-cued words than remember-cued words, but this effect disappeared for participants exposed to the list method (see MacLeod, 1998 for a review).

In their seminal paper, Basden, Basden, and Gargano (1993) compared the item and list methods directly to explore possible underlying mechanisms responsible for the discrepant findings (see MacLeod, 1999 for a replication). They argued that selective rehearsal was responsible for DF effects in both recall and recognition tasks using the item method, but this was an insufficient explanation for response patterns generated using the list method. When participants receive the forget/remember cue immediately after each word, it is feasible to suspend rehearsal of the forget-cued items shortly after presentation of the item. Thus, remember-cued words receive more rehearsal relative to forget-cued words and DF effects for both recall and recognition tasks are said to result from differential encoding. Using the list method, the forget/remember cue does not appear until after presentation of the entire list, making selective rehearsal more difficult. Thus, DF effects in recall performance using the list method are said to result from inhibition: When participants are first instructed to forget List 1 and then presented with a list of items they are instructed to remember (i.e., List 2), a process is initiated that inhibits retrieval of the forget-cued items (E. L. Bjork & R. A. Bjork, 1996; 2003). This inhibition prevents forget-cued items from interfering with recall of remember-cued items, leading to DF effects in recall performance. However, given that participants recognize forget- and remember-cued items equally well in list method studies, inhibition does not appear to influence the strength of the memory for forget-cued items. Indeed, some argue that recognition questions may provide a release from inhibition because participants are presented with the items, and so no item generation is needed (e.g., for more detailed discussions of inhibition see R. A. Bjork, 1989; Harnishfeger & Pope, 1996).

For the most part, consensus in the DF literature is that selective rehearsal best explains response patterns obtained using the item method and inhibition best explains response patterns obtained using the list method (e.g., Basden & Basden, 1998; R. A. Bjork, 1989; Basden et al., 1993). However, recent list method studies call this latter point for consensus into question. MacLeod, Dodd, Sheard, Wilson, and Bibi (2003) reviewed findings from several areas of attention and memory research (e.g., negative priming, retrieval-induced forgetting, and DF), and presented arguments *in opposition to inhibition*. They began their critique of the role of inhibition in DF by summarizing a series of studies (Sheard, Dodd, Wilson, & MacLeod, 2002, as cited in MacLeod et al.) that used procedures similar to Basden and Basden (1998). There were three experimental conditions: (1) delay-warning (prior to a delay of several minutes, participants were informed that the recall task would include both R-cued and F-cued items), (2) delay-no warning, and (3) standard no delay-no warning condition. Compared to participants in the standard DF condition, participants in the delay-warning condition evinced smaller DF effects and participants in the delay-no warning condition showed larger DF effects. They argued that selective rehearsal accounted for their findings.

Specifically, during the delay, participants who received a warning about the upcoming recall task rehearsed F-cued items more than R-cued items because participants anticipated F-cued items would be most difficult to recall. This differential rehearsal would lead to a reduction in the DF effect. Contrast this with the increased DF effect with participants in the no warning condition who, presumably, spend more time rehearsing R-cued words relative to F-cued words during the delay.

Sheard and colleagues (2002, as cited in MacLeod et al., 2003) replicated the aforementioned findings in a follow-up study, wherein rehearsal was controlled. Specifically, participants in one condition were discouraged from rehearsing (i.e., by filling the delay between presentation and test with an effortful task) and participants in the other condition were encouraged to rehearse (i.e., prior to the delay between presentation and test participants were informed of a financial incentive to recall as many words as possible). As in the first study, half of the participants in each rehearsal condition were warned about the content of the upcoming recall test and half of the participants were not. As expected, performance of participants encouraged to rehearse matched those from the first study: Increased DF effects in the delay-no warning condition and decreased DF effects in the delay-warning condition relative to the standard no delay-no warning condition. This pattern of findings was not true for participants discouraged from rehearsal. Instead, they demonstrated equivalent DF in the delay-warning and delay-no warning conditions, both of which were smaller than DF effects found for participants in the standard no delay-no warning condition. If inhibition was responsible for list method findings then encouraging or discouraging rehearsal should not influence the size of the DF effect. The researchers preferred the parsimony of a selective rehearsal explanation for both item method and list method DF effects to having separate explanations for two similar procedures. Granted, these studies support a selective rehearsal account for list method DF findings under conditions of delayed recall; MacLeod and colleagues agreed additional work was necessary to determine whether this may also be the case for immediate recall.

Sahakyan and colleagues (e.g., 2002, 2003) have proposed an entirely different explanation for list method findings, namely that two separate mechanisms underlie costs and benefits. Support for this two-factor approach can be found in several studies documenting costs with no benefits (e.g., Conway, Harries, Noyes, Racsma'ny, & Frankish, 2000), or benefits with no costs (e.g., Macrae, Bodenhausen, Milne, & Ford, 1997). Typically, scholars have interpreted costs arising from inhibition of the forget items, and benefits arising from the subsequent reduced proactive interference due to fewer List 1 items in memory (e.g., E. L. Bjork & R. A. Bjork, 1996; R. A. Bjork, 1989). However, Sahakyan and colleagues proposed a two-factor model with different mechanisms for DF costs and benefits, whereby costs result from contextual change brought on by the forget cue and benefits arise when participants adopt more effective study strategies when encoding List 2 (e.g., Sahakyan & Delaney, 2005). According to the *context change hypothesis,* forget-cue costs occur because the List 1 study context (forget these items) differs from the test context (recall items from both lists). In support of this hypothesis, Sahakyan and Kelley (2002) obtained findings similar to DF costs after merely inducing a context change between lists for remember-cue only participants (i.e., before presentation of List 2 items, participants were asked to generate thoughts about a completely irrelevant topic). Further evidence for the *context change hypothesis* was found in Sahakyan (2004). She found costs of DF after modifying the DF list method protocol to include three word lists instead of two. Three DF conditions resulted: FRR (i.e., forget List 1 and remember Lists 2 and 3), RFR (i.e., remember List 1, forget List 2, and remember List 3), and RRR (i.e., remember all three lists). The effects Sahakyan used to support the *context change hypothesis* relate most to those participants in the RFR condition, wherein instructing participants to forget List 2 words impaired access to List 1 words as well. Sahakyan referred to this as *indirect* costs of the forget cue. In other words, participants in both the FRR and RFR conditions recalled fewer List 1 words relative to participants in the RRR condition. Sahakyan argued that the forget cue causes participants to establish a new context; thus, for participants in the RFR condition, the recall context more closely resembled the study context for List 3 than the study context for List 1. Findings from Experiment 2 strengthened her argument. Even when lists contained categorically related words (e.g., List 1 words were vegetables, List 2 words were animals, and List 3 words were fruits), participants in the FRR and RFR conditions recalled fewer List 1 words than participants in the RRR condition.

Upon examination of retrospective verbal reports from participants in Sahakyan and Kelley (2002), it was evident that the *context change hypothesis* was not sufficient to explain DF benefits. Interestingly, compared to participants instructed to remember List 1, participants instructed to forget List 1 were more likely to switch from using a shallow encoding strategy (e.g., rehearsal) to a deeper encoding strategy (e.g., creating a story using all the items on the list) for List 2. In two experiments, Sahakyan and Delaney (2003) found that controlling for encoding strategy (i.e., participants were instructed when to use deep/shallow strategies) eliminated DF benefits. Sahakyan, Delaney, and

Kelley (2004) were able to eliminate DF benefits as well by prompting self-evaluation of study strategy in all participants. They encouraged participants to evaluate their study strategies, before issuing the DF instructions, by asking participants to recall each list immediately after learning (Exp. 1) and by making predictions about how many words from List 1 they were likely to recall (Exp. 2). Benefits were not observed in either experiment for participants prompted to evaluate their study strategies, but typical DF benefits were observed for control participants in the non-evaluation conditions. In other words, when participants in the RR condition were encouraged to evaluate their study strategy midway through learning, they were able to recall List 2 as well as participants in the FR condition.

In summary, numerous studies have examined the influence of instructing an adult to forget words on later recall and recognition of those words. Using either item-cueing or list-cueing leads to difficulty recalling forget-cued words relative to remember-cued words; however, recognition of forget-cued words is impaired only under item method conditions. This pattern of findings has led to ongoing debate regarding what mechanism(s) is/are responsible for the DF effect. Consensus in the literature suggests that item method findings (i.e., DF effects for both recall and recognition tasks) result from selective rehearsal: Participants suspend rehearsal of forget-cued items upon presentation of the forget cue, but continue to rehearse remember-cued items. The debate surrounding underlying mechanisms has focused on list method findings (i.e., DF effects for recall but not recognition). The most common interpretation identifies inhibition: Participants encode both forget-cued and remember-cued words because the cue follows presentation of an entire list of words, thus inhibition impairs *recall* of forget-cued words and *recognition* provides a release from inhibition. However, findings from recent studies are cause for a re-evaluation of the inhibition proposal. MacLeod and colleagues (2003) believe selective rehearsal can account for both item method and list method findings, whereas Sahakyan and colleagues (e.g., 2002, 2003, 2004) propose two entirely different mechanisms for list method costs (i.e., context change) and benefits (i.e., encoding strategy). Regardless of what underlying mechanism is at work, the DF literature has demonstrated that adults are able to direct their forgetting of words. Indeed, these "costs" are the focus of proposals that DF may, at least in part, explain delayed disclosure of childhood sexual abuse. Certainly, motivation to forget these troubling experiences far outweighs that felt when participating in a laboratory experiment of word list memory. Although this argument is compelling, using the basic DF protocol with children and in applied settings is imperative.

### *Directed Forgetting and Children*

Although the majority of DF studies have involved adult participants, researchers began to examine children's ability to adhere to DF instructions early on (e.g., Posnansky, 1976). Studies with children are fewer in number than those conducted with adults, but developmental investigations have informed the debate regarding underlying mechanism(s) of the DF effect. Generally, these studies consistently demonstrate developmental differences in DF effects (see Wilson & Kipp, 1998 for a review). Noteworthy is that children exhibit difficulty recalling forget-cued items at different ages depending on method of cue presentation. On the one hand, DF effects in item method studies emerge in the early elementary school years (i.e., grades one, two, three), but these effects are smaller relative to adult performance (e.g., Lehman & Bovasso, 1993;

Posnansky). On the other hand, in list method studies, children do not exhibit significant DF effects until later in the elementary school years (e.g., grade five; Bray, Justice, & Zahm, 1983). Differences in onset of DF effects as a function of method of cue presentation have been used as evidence for the role of different mechanisms in item and list method findings. As in the adult literature, the oft-cited proposal in developmental DF studies is that selective rehearsal accounts for item method findings and inhibition accounts for list method findings.

Researchers have proposed that developmental differences found in item method studies can be explained by children's developing ability to suspend rehearsal of forget-cued items and to encode remember-cued items more elaborately (e.g., Lehman, McKinley-Pace, Leonard, Thompson, & Johns, 2001). Lehman and Bovasso (1993) created a novel and child-friendly version of item-cueing, in which they read children 7, 9, and 11 years of age a story about a bee searching for honey. They informed children that places the bee did not find honey were to be forgotten (i.e., these locations were followed by a page containing a large, red X) and places the bee did find honey were to be remembered (i.e., these locations were followed by a page containing a picture of a honey pot). Participants completed an initial recall task, after which they were presented with all items again as remember-cued only, and then completed a second recall task. They argued that, if retrieval inhibition (as opposed to rehearsal) was responsible for their DF findings, then participants should exhibit better recall of previously unreported forget-cued locations on trial two than previously unreported remember-cued locations. The logic underlying this hypothesis was twofold; whereas remember-cued items benefit from a second rehearsal (i.e., presentation of all items a second time as remember-cued) only, forget-cued items benefit from a second rehearsal and a release from inhibition. All participants recalled previously unreported forget-cued locations and remember-cued locations at similar rates during the second recall task, and the authors concluded that selective rehearsal was the most supportable explanation for their findings.

Consistent with the adult literature, scholars who conduct developmental DF research continue to advocate the role of inhibition in list method DF effects (e.g., Wilson & Kipp, 1998; Lehman, McKinley-Pace, Wilson, Slavsky, & Woodson, 1997; but see Lehman, Srokowski, Hall, Renkey, & Cruz, 2003 who argue that inhibition may also play a role in their item method findings). Harnishfeger and Pope (1996) used list-cueing with child and adult participants and replicated typical list method DF effects for older children (grade five) and adults: Items from forget-cued lists were recalled less often than items from remember-cued lists, but recognition accuracy was unaffected by cue. Children in grade three showed some forgetting as a result of the forget-cue, but less effectively than older children and adults. The youngest participants (grade one) did not exhibit any DF effects, recalling items from forget-cued lists as often as items from remember-cued lists. Bray and colleagues (1983) found a similar pattern of results, such that children in grade five demonstrated the most sophisticated and adult-like DF effects. For those who subscribe to an inhibition account of list method effects, these findings may be used to suggest that children's developing inhibitory skills help to explain developmental differences in the DF effect.

Harnishfeger and Pope (1996) cautioned against making claims of the development of inhibition without also considering the way in which inhibition is defined and measured. In addition to the retrieval inhibition (i.e., the inability to generate forget-cued words, and most common type of inhibition referenced in the DF literature) findings described above, Harnishfeger and Pope also examined something they referred to as response inhibition (i.e., participants withhold overt recall of forget-cued words). They randomly assigned participants to one of three recall conditions: Remember-All (instructed to remember both lists and asked to recall all list items), Forget-Only (instructed to forget List 1, remember List 2, and asked to recall List 2 items only), and Forget-All (instructed to forget List 1, remember List 2, and asked to recall all list items). They hypothesized that participants with better response inhibition would be able to withhold a remembered response, as evidenced by greater List 1 recall for participants in the Forget-All condition than participants in the Forget-Only condition. In other words, participants intentionally withhold the production of remembered forget-cued items. Only adults did this; none of the child participants (grades one, three, and five) exhibited signs of response inhibition. Wilson, Kipp, and Daniels (2003) expanded upon this methodology by manipulating category relatedness of word lists, as the task of inhibiting a response is easier when items are related than when they are not related. Retrieval inhibition patterns of recall began to appear by six years of age when list items were related, but similar to Harnishfeger and Pope, retrieval inhibition patterns of responding were not found until eight years of age when list items were unrelated. Children as young as eight exhibited signs of response inhibition when list items belonged to the same semantic category, but neither the six-year-olds nor the eight-year-olds evinced response inhibition when list items were unrelated. These findings suggest that children can demonstrate inhibitory skills at different ages, depending on task difficulty.

In general, DF studies with children have shadowed the adult literature by evaluating children's recall and recognition of words in a list or story, and by focusing on mechanisms responsible for DF effects. Indeed, as with DF in adults, the developmental literature posits that selective rehearsal best explains item method findings and inhibition accounts for list method findings. There are a limited number of DF studies with children, but they have provided invaluable insight into the development of DF. Overall, these studies demonstrate that it is more difficult to obtain DF effects with children than with adults, using either item- or list-cueing procedures, as children often recall forget-cued words in spite of the DF instructions. Nevertheless, children evidence DF earlier when using item-cueing than list-cueing. There are, however, many unexplored avenues to further our understanding of the development of DF. For example, exploring the application of DF instructions to children's autobiographical memory reports is critical to justify claims that DF may result from attempts to avoid rehearsal of CSA memories. Coupled with findings from laboratory DF studies, namely that forget-cued details often surface in younger children's recall (i.e., younger children do not exhibit DF effects), it may be especially important to conduct applied DF research with young children to determine the extent to which DF instructions can affect later recall of an event.

### *Directed Forgetting and Autobiographical Memory*

In general, studies using the DF protocol have focused exclusively on memory for words in a list, sentence, or story. To bolster support for the notion that avoiding traumatic memories may lead to DF effects, recent years have marked the emergence of DF studies using different stimuli like emotional words (see Geraerts & McNally, 2008 for a review) and autobiographical memories. Unfortunately, the least researched application of the DF protocol may be the most relevant to the forensic context to which basic laboratory DF research is applied. Based on findings from the literature, primarily that recall of forget-cued items is difficult, scholars have argued that DF may be responsible for some autobiographical memory loss. Anderson and Green (2001) proposed that individuals try to avoid cues that remind them of unwanted memories, and that forgetting increases with number of attempts to avoid these memories. They argued further that memories of traumatic experiences, like those associated with CSA, might be especially prone to forgetting because of increased motivation to avoid thoughts of the abuse and many opportunities for avoidance during delayed disclosure. However, autobiographical events differ from word lists in several important ways that may change the potential application of DF to autobiographical memory. For example, most autobiographical events are more personally relevant and salient than word lists. Similarly, events are rich with sensory information (i.e., scent, touch) and detail; a simple instruction to forget an event, or parts of it, may not be effective given the complexity of autobiographical memories. At the very least, differences between memory for word lists and memory for events illuminate the importance of conducting applied research on the DF effect.

Applications of the DF protocol to memory for events are limited; only two publications exist and both included adult samples. Joslyn and Oakes (2005) conducted two experiments wherein participants were instructed to record 10 events per week for two weeks in a diary. After the first week of diary-keeping, participants were invited back to the lab and issued the DF instructions. Half of the participants were informed that Week 1 events were not part of the memory study, and that an effort should be made to forget Week 1 events to facilitate memory for Week 2 events. The remaining participants were informed that Week 1 events were part of the memory study and attempts should be made to remember these events in addition to the upcoming Week 2 events. Cost/benefit analyses for both experiments demonstrated costs of the forget cue: Compared to participants in the remember condition, participants in the forget condition recalled a smaller proportion of Week 1 events. Benefits of the forget cue (i.e., increased recall of Week 2 events when Week 1 events were forget-cued) were not found. They argued failure to find benefits might have resulted for one of two reasons. First, although researchers assert that reduced proactive interference facilitates benefits of the forget cue, this may have been negligible for participants given the likely large number of autobiographical events experienced in Week 1 that were not recorded and so not the subject of a forget instruction. Second, the cover story provided at the outset of the experiment (i.e., you may or may not participate in a memory study based on events that transpire in Week 1, but certainly for those that occur in Week 2) may have encouraged all participants to switch to more effective study strategies for Week 2 events. After Week 1, the experimenters re-iterated to all participants that Week 2 events definitely would be subject to memory recall. Thus, it was difficult to ascertain whether the experiments' stimuli or procedures contributed to the lack of benefits.

In a series of six experiments, Barnier and colleagues (2007) demonstrated that list method DF instructions could influence recall of autobiographical memories as well. Participants in these studies were given two lists of words and asked to generate a unique autobiographical memory related to each word in the list. After presentation of all items, participants were asked to recall both the word and description of the associated memory. Costs associated with the forget cue were found for all six experiments: Participants recalled fewer List 1 autobiographical memories when they were forget-cued than when they were remember-cued. These costs were observed regardless of whether participants rated emotionality of the memories as positive, neutral, or negative. However, several factors modified the influence of forget instructions on recall costs. For example, similar to recognition task performance using the list method, cued recall questions eliminated the DF effect. In addition, costs were reduced when participants were asked to recall events from a similar period in their lives. As in Joslyn and Oakes (2005), the forget cue did not facilitate recall of List 2 memories (i.e., benefits). Barnier and colleagues argued their reliable and consistent findings across six experiments were a demonstration of the successful application of DF to autobiographical memory reports.

These novel procedures mark the beginning of an important area of research, as findings from the aforementioned two papers suggest that DF instructions can influence later recall of autobiographical events. Specifically, these studies demonstrate that instructing an individual to forget autobiographical events leads to impoverished recall of those events. Importantly, forget-cue benefits were not found. This is particularly interesting given the forensic context to which these data can be applied, as DF is often used to explain failures to report adverse experiences. Given the paucity of research, numerous unexplored areas of study remain. The present research aimed to extend this line of inquiry by examining the influence of DF instructions on both child and adult participants' recall of a personally experienced event.

### *Executive Functioning*

Although selective rehearsal and inhibition provide insight into why developmental differences in obtaining DF effects exist, there has been little investigation into possible individual differences that may account for children's performance on DF tasks. Recently, researchers have documented the importance of exploring the relationship between individual differences, like executive functioning, and children's memory (e.g., Roberts & Powell, 2005). Executive functioning (EF) is an umbrella term used to describe several related cognitive skills (e.g., self-regulation, planning, cognitive flexibility, inhibition, selective attention) and research demonstrates shared relationships between EF and theory of mind (e.g., Carlson & Moses, 2001), source monitoring (e.g., Ruffman, Rustin, Garnham, & Parkin, 2001), suggestibility (e.g., Melinder, Endestad, & Magnussen, 2006), and academic success (e.g., van der Sluis, de Jong, & van der Leij, 2007; Mazzocco & Kover, 2006). Although the literature on EF is vast, at least two problems plague researchers: defining EF and measuring EF.

Definitions of EF differ widely across studies and evidence for the number and type of cognitive skills affiliated with executive functions is equivocal. A common approach to defining EF involves creating a comprehensive battery of tasks that tap into various executive functions and using factor analysis to uncover underlying structural elements. Generally, these studies find three or four factors; a finding often used to suggest that EF is not unidimensional. Three factors of EF commonly described in the literature are shifting, updating, and inhibitory control (e.g., Brocki & Bohlin, 2004; Miyake et al., 2000). Shifting is the ability to switch between sets, tasks, or strategies efficiently. Updating refers to the ability to evaluate incoming information for relevance to the task at hand and then revising the content of working memory by replacing old (i.e., irrelevant) information with new (i.e., relevant) information (Morris & Jones, 1990). Inhibitory control involves suppressing dominant responses in favour of more goal-appropriate responses. All three of these executive functions may be involved in how successfully an individual adheres to DF instructions. In fact, inhibition is the very mechanism thought by many to be responsible for list method findings. After receiving an instruction to forget items, specifically that the items need not be recalled later, an individual may revise the content of his/her working memory to include only those remember-cued items (i.e., updating). Lastly, how effectively and efficiently an individual is able to switch between remember-cued and forget-cued items may influence whether a DF effect occurs (i.e., shifting). Indeed, recall that Sahakyan and colleagues (e.g., 2003, 2005) found DF benefits when participants were encouraged to switch to more effective study strategies between lists.

The second problem plaguing EF researchers relates to variability and difficulty in the measurement of executive functions, referred to as task impurity. Manifestation of executive functions is task-dependent, but this leads to problems in interpretation as tasks may require multiple executive functions or non-executive cognitive skills such as verbal ability or motor speed for completion. In addition, many tasks measuring EF call on working memory (e.g., Tower of London; Sikora, Haley, Edwards, & Butler, 2002). To address this, many scholars encourage the inclusion of multiple tasks to measure the same executive function (e.g., van der Sluis et al., 2007). Unfortunately, due to time constraints, this was not possible in the current study. However, in addition to one measure each for shifting, updating, and inhibitory control, a measure of working memory was included. The current study aimed to expand recent discussions into the role of individual differences on children's memory by investigating whether EF accounts for some of the developmental differences in children's reports of a personally experienced event containing DF instructions.

### *Repeated Interviews*

Participants in DF experiments generally complete one memory interview, consisting of free recall and recognition questions, following presentation of the items. In many cases, the recognition task reduces or eliminates the DF effect. The applicability of these procedures to the forensic context is limited for several reasons: First, repeated interviewing is common and scholars have speculated that a witness may be subjected to upwards of 12 interviews throughout the course of an investigation (e.g., Bala, 1999; Ceci & Bruck, 1993; Whitcomb, 1992); second, forensic interviewers often do not have access to information about the offense to facilitate creation of a recognition task; third, best practice models of forensic interviewing advocate the use of open-ended questioning (see Milne & Bull, 1999 for a review), and recognition questions may be considered leading or suggestive, thereby challenging admissibility of the evidence. By all of these indicators, alternative techniques for reducing omissions are necessary. Numerous studies have investigated the influence of repeated interviewing on report content and found several benefits to this practice. Repeated interviews can provide an inoculation against forgetting (e.g., Baker-Ward, Hess, & Flannagan, 1990), encourage recall of previously unreported information (e.g, Salmon & Pipe, 1997), and increase accurate recall across retrieval attempts (e.g., La Rooy, Pipe, & Murray, 2005). Thus, multiple recall attempts may provide a viable option for forensic interviewers attempting to elicit comprehensive reports from witnesses.

It should be noted that there is also a substantial literature on the consistency of children's reports in response to repeated questioning *within* a single interview (for reviews see Ceci & Bruck, 1993; Fivush, & Schwarzmueller, 1995; Poole & White, 1995). Overall, the findings mirror those of the repeated interviewing studies. Specifically, when repeated questions are open-ended, inconsistencies are uncommon, participants provide new information in later trials, and the amount of inaccurate information does not increase across trials (e.g., Hershkowitz & Terner, 2007; Memon & Vartoukian, 1996; Poole & White, 1991). However, when repeated questions are closed-ended (i.e., recognition questions), suggestive, or misleading, interviewees are more likely to change their responses (e.g., Brady, Poole, Warren, & Jones, 1999). Participants in the current study were not subjected to repeated questions within a single interview; thus, the literature review focuses on the effects of repeating questions across multiple interviews.

Often, child witnesses are questioned repeatedly throughout the course of a forensic investigation in a variety of contexts (informally with family members and formally with social workers, police officers, and lawyers) for a variety of purposes (to collect additional evidence, clarify previous statements, or even to confuse or mislead the child). Typically, studies that report benefits to multiple interviews use neutral open-ended questioning (e.g., "Tell me everything you remember from that time") and include intervening interviews shortly after the event. In their seminal paper, Dent and Stephenson (1979) demonstrated that 10- and 11-year-old children's long-term retention of event information was improved by participation in intervening interviews. Specifically, children interviewed four times (immediately, one day, two days, and two weeks after the event) and children interviewed once (two weeks after the event) before the final interview two months after the event provided more correct information than children interviewed after two months only. These findings have been replicated several times, suggesting that early interviews assist with consolidation of memory for an event (e.g., Baker-Ward et al., 1990; Hudson, 1990; Warren & Lane, 1995).

Repeated interviewing does not come without its costs, however, particularly when the questions asked across interviews are misleading or suggestive (e.g., Bjorklund, Bjorklund, Brown, & Cassel, 1998; Ceci & Bruck, 1995; Leichtman & Ceci, 1995; Quas & Schaaf, 2002). In addition, Poole and White (1991) demonstrated that report accuracy and consistency change in response to repeated interviewing when interviews contain closed-ended questions. Comparing responses after a one-week delay with responses immediately after the event, they found that 4-year-old children were more likely to change their responses to yes/no questions than older children (6- and 8-year-olds) and adults. Furthermore, all participants provided more "I don't know" responses to the yes/no questions in the second interview than in the first interview, suggesting that participants were providing less information over time. Poole and White (1993) conducted a follow-up study two years later and compared participants' responses during the follow-up interview to their responses one week after the event. They found that children provided consistent responses to yes/no questions only half of the time. Interestingly, the majority of children changed their responses from accurately reporting bad behavior exhibited by the male intruder to inaccurately denying his transgressions (i.e., that he asked for the pen nicely and did not hurt his colleague). Poole and White (1993) suggested this might have occurred because the delay to test weakened children's memory for the altercation. Overall, the risk for inaccuracy and inconsistency across repeated interviews appears to increase when interviews contain leading, suggestive, or closed-ended questions.

Studies with legal professionals and laypersons have uncovered a pervasive notion that consistency is one of the strongest predictors of perceived accuracy of an eyewitness account (e.g., Brewer, Potter, Fisher, Bond, & Luszcz, 1999; Potter & Brewer, 1999). Thus, inconsistencies can undermine witness credibility because perceptions of witness accuracy will be compromised. Indeed, compared to inconsistent witnesses, previous research shows that consistent witnesses receive higher ratings on witness effectiveness (e.g., Goodman, Goldings, & Haith, 1984; Goodman et al., 1998), defendant culpability (e.g., Brewer & Burke, 2002; Brewer & Hupfeld, 2004), and prosecution case credibility (e.g., Semmler & Brewer, 2002). Although consistency appears to influence *perceived* accuracy, the relationship between consistency and *actual* accuracy is moderate at best (e.g., Brewer et al., 1999, Exp. 2; Fisher & Cutler, 1995). That report consistency has considerable weight in an assessment of the veracity of a statement presents two unsettling consequences for child witnesses who omit information from their memory reports. On the one hand, inconsistent accounts of abuse experiences may unfairly diminish child witness credibility. On the other hand, consistent omissions of abuse details place children at risk of continued abuse.

Given the potential benefits to repeated open-ended questioning (i.e., inoculation against forgetting, reminiscence, and hypermnesia); multiple recall attempts may provide the support necessary to reduce the likelihood that children omit information consistently, but this may depend on the contribution of inhibition and/or selective rehearsal in generating DF effects. For example, repeated recall attempts are unlikely to aid recall of previously unreported information in cases where selective rehearsal led to omissions.

However, repeated recall attempts may provide a release from inhibition. In other words, this study examined whether multiple recall attempts can reduce lasting effects of DF instructions on reports of a personally experienced event.

### *The Present Research*

The current study explored the influence of DF instructions on memory for a personally experienced event. Specifically, it focussed on developmental differences in obtaining DF effects, whether developing executive functions correspond to DF findings, and if repeated recall attempts eliminate/reduce the influence of DF instructions on children's memory reports. School-aged children and undergraduate students participated in two types of sessions: an activity session and two interviews. Participants completed four activities during the activity session. For half of the participants, forget instructions were issued immediately after the first two activities (i.e., List 1) and remember instructions were issued after the last two activities (i.e., List 2). The remaining participants were instructed to remember the activities, once after they completed List 1 activities and again after they completed List 2 activities. Six days later, participants completed two executive functioning tasks. In addition, participants in the repeated recall condition were asked to describe what occurred during the activity session. One day later, all participants recalled the activity session, answered recognition questions about specific details contained within each activity, and completed the remaining executive functioning tasks.

It was expected that DF instructions would have their greatest impact on adult participants, who would demonstrate both DF costs (i.e., recalling fewer List 1 details when they were forget-cued than when they were remember-cued) and benefits (i.e., participants in the FR condition recalling more List 2 details than participants in the RR condition). These significant DF effects were also expected for older children (grade five students) but to a lesser degree than what was expected of adults. Reports from grade three students may include some signs of following the DF instructions, but it was unlikely that DF instructions would influence the content of grade one students' free recall reports. The inclusion of individual differences measures in the current study was exploratory, but it was expected that EF might account for some of the developmental differences in DF effects. Specifically, shifting, updating, working memory and inhibitory control skills may be related to how well participants direct their forgetting of the play session. As is typically found in DF studies using the list method, DF effects were not expected for recognition responses for all participants. Recall that many scholars argue the absence of DF effects in recognition is a demonstration of release from retrieval inhibition. Given past research that suggests repeated interviews can lead to increased recall across retrieval attempts (reminiscence; e.g., Salmon & Pipe, 1997), multiple recall attempts may facilitate a release from inhibition. Thus, if retrieval inhibition is responsible for DF findings using the list method, then participation in the intervening recall session should lead to an increase in recall of forget-cued details relative to remember-cued details in the subsequent memory interview one day later. However, if inhibition does not account for list method DF effects, then repeated recall attempts should not facilitate recall of previously unreported forget-cued details. Failure to find increased recall of forget-cued items across interviews may be indicative of selective rehearsal, as an intervening recall attempt should have no influence on the effects of differential rehearsal.

# METHOD

## *Participants*

There were 246 participants in this study, belonging to one of four age groups: grade one students ($n$ = 64, $M$age = 6.90 years, $SD$ = 0.36 years), grade three students ($n$ = 58, $M$age = 8.90 years, $SD$ = 0.36 years), grade five students ($n$ = 60, $M$age = 10.93 years, $SD$ = 0.35 years), and university students ($n$ = 64, $M$age = 19.80 years, $SD$ = 2.17 years). Child participants were recruited through Catholic elementary schools in the greater Vancouver area. Principals in the participating school district were contacted and asked to grant permission to conduct the study, after which teachers and parents were approached to contact individual students. Only those children who received parental permission to participate in the study completed the sessions (approximately 50% of students from each classroom received parental permission). Participating classrooms received a $50 gift certificate to an educational supplies store and each child from participating classrooms received a pen or pencil of their choice. Adult participants were students from the Simon Fraser University community and recruited via the Research Participation System. Adults received course credit for participating in the study.

## *Materials*

Participants completed one task each for shifting, updating, working memory, and inhibitory control.

**Shifting.** Recall that shifting is the ability to switch between sets, tasks, or strategies efficiently. The Making Trails (e.g., Espy & Cwik, 2004; Reitan, 1992) task is similar to connect-the-dots tasks found in children's activity books and often used as a measure of shifting. Typically, participants complete two trails. For Trail A, participants connect 15 randomly distributed numbered circles in numeric order (i.e., 1-2-3-4-etc.) as quickly as possible. For Trail B, participants connect numbered (1-13) and lettered (A-L) circles in alternating order (i.e., 1-A-2-B-3-C-etc.) as quickly as possible. In total, Trail B consists of 25 circles and participants must shift between connecting numbers and letters to complete it successfully. The experimenter instructed participants to correct errors before moving on to the next circle and recorded the total time required to complete the task (measured with a stopwatch) on a separate sheet of paper. The Making Trails task has been widely used as a measure of shifting with participants as young as seven years of age (e.g., Im-Bolter, Johnson, & Pascual-Leone, 2006). Although many of the grade one students in the current study were younger than seven years of age, there is no time limit for completing each trail. Thus, that the youngest participants took longer to complete the trails than other participants does not reduce the task's ability to measure shifting. Instead, it is more important that participants take less time to complete Trail A than Trail B (see results section for these findings).

**Updating.** Updating refers to revising the content of working memory by replacing old (i.e., irrelevant) information with new (i.e., relevant) information (Morris & Jones, 1990). The Keep Track task required participants to remember the last category exemplar from a series of pictures. There were five different categories: weather (sunny, rainy, cloudy, or snowy), shapes (triangle, square, circle, or rectangle), sports (soccer, basketball, hockey, or baseball), fruits (apple, orange, banana, or grape), and animals (tiger, lion, dog, or cat). Category exemplars were obtained from normative data with children in Price and Connolly (2006). Before beginning any test trials, the experimenter showed participants a legend of all four exemplars for each of the five categories and named each picture. This ensured that all participants understood what the pictures depicted. Each trial began with the experimenter naming from which categories participants would be asked to recall. Two trials required participants to recall the last exemplar from three categories and two trials required participants to recall the last exemplar from four categories. At the end of a trial, participants were asked to name the last picture from each of the target categories. For example, correct responses to a trial consisting of "triangle, cat, hockey, cloudy, circle, banana, soccer, orange, sunny, square" when target categories were shapes, animals, and fruits would be "square, cat, and orange". Each trial contained ten pictures; a random number generator created the order in which pictures were presented, with the restriction that each picture was presented at least once throughout the four trials. The experimenter recorded number of correct responses on a separate scoring sheet. Researchers have performed this task with participants as young as eight years of age (e.g., van der Sluis et al., 2007). To accommodate younger participants in the current study, all participants completed a practice trial with two target categories only and reviewed their answers with the experimenter. If the participant did not correctly recall both category exemplars, they completed a second practice trial. Most participants (78.0%) did not require the second practice trial.

**Working memory.** Working memory was assessed by administering the Digit Backwards task from the Test of Memory and Learning (TOMAL; Reynolds & Voress, 2007). In this task, the experimenter read a series of digits and asked participants to repeat the series in the opposite order (e.g., if the experimenter read digits 4 and 6 the correct answer was 6 and 4). Trials increased in difficulty, ranging from 2-9 digits. There were 13 possible trials: two trials with two digits, two trials with three digits, two trials with four digits, two trials with five digits, two trials with six digits, one trial with seven digits, one trial with eight digits, and one trial with nine digits. The experimenter ended the task if a participant repeated only three or fewer digits in the correct position in a single trial. The raw score was the total number of digits recalled in the correct position. A proportion correct score was created for each participant, wherein the total number of digits participants repeated in the correct position was divided by the total number of digits presented by the experimenter across all trials.

**Inhibitory control.** Inhibitory control involves suppressing dominant responses in favour of more goal-appropriate responses. The primary objective of the Stroop Color-Word test (Stroop, 1935) is to name the colour of a set of stimuli as quickly as possible, often used as a measure of inhibitory control in the EF literature. Administration involved participants reading three separate pages of items: word page, colour page, and colour-word page (Golden & Freshwater, 2002; Golden, Freshwater, & Golden, 2003). For the word page, participants read as many words (i.e., RED, GREEN, BLUE) as they could in the time allotted. Words were printed in black ink, and scores on this task provided a baseline measure of participants' reading proficiency. The colour page consisted of columns of "XXXX"s printed in different colours (i.e., RED, GREEN, BLUE), and participants identified the colour of ink in which each set of X's were printed. The colour-word page was similar to the word page, with the exception that words were printed in ink incongruent with the meaning of the word (e.g., the word GREEN printed in RED ink). Participants identified ink colour rather than merely reading the words. Participants corrected errors before moving on to the next item, and the number of items identified correctly was recorded for each condition to create raw scores. To control for individual reading and colour naming speed, a score was created that involved subtracting the number of items read on the colour page from the number of items read on the colour-word page (i.e., inhibitory control = CW – C). Thus, higher scores reflect greater inhibitory control. This score is ideal for younger participants with developing reading proficiency and was the suggested measure of inhibitory control in the administration manual for children (Golden et al.).

### *Design and Procedure*

This was a 4 (Age: grade one, grade three, grade five, adult) × 2 (Directed Forgetting: FR, RR) × 2 (Interview: single, intervening) between-subjects design. Participants were randomly assigned to each of the 16 experimental conditions (*n*'s ranged from 12-16 for each cell). Participants completed three sessions with three different experimenters lasting approximately 20 minutes each: an activity session and a memory interview in two parts.

**Activity sessions.** Participants, in groups of up to three, participated in an activity session consisting of four separate activities: paper folding, magic trick, sticker colouring, and clay molding. Each activity contained five critical details, for a total of 20 details (see Table 1 for a complete list of details). The activity session leader issued DF instructions twice during the session: once after completing the first two activities (i.e., halfway through the play session) and again after completing the second two activities (i.e., end of the play session). The activity session leader instructed half of the participants to forget the first two activities and the remaining participants to remember the first two activities. All participants were instructed to remember the third and fourth activities. For example, participants assigned to the FR condition received the following instruction after completing the first two activities: *"Try to forget the activities we just did. They were just practice so we could get to know each other better. You will not need to think of them again later."* After completing the second two activities, the activity session leader instructed participants to, *"Focus on these activities and try to remember them because you will be asked about them later."* Participants were randomly assigned to one of two different activity orders to ensure the forget/remember cues were assigned to all four different activities. One of two trained experimenters, each with extensive experience working with children, conducted all activity sessions. The activity session leader wore a silver bowtie and introduced the session as "Bowtie Playtime", which tagged the participants' memory for the session during the interviews. Activity sessions were conducted outside of the classroom with child participants and in a laboratory room decorated like an elementary school classroom with adult participants. Completing the activities in groups allowed child participants to feel more comfortable, as they were working with their classmates, and minimized the embarrassment of completing childish activities for adult participants. As a whole, all participants appeared to enjoy the activities and were engaged throughout the session.

**Memory interviews.** All participants were interviewed six days after the activity session and again one day later, but the content of these interviews changed as a function of interview condition. Six days after the activity session, all participants completed two of the four executive functioning tasks (two random orders of tasks were created). In addition, the interviewer asked participants in the intervening interview condition to recall the activity session. Free recall was conducted in the same manner for both interviews, and is described below (see Appendix A for an example of the first interview). One day later, all participants completed an identical memory interview consisting of free recall, recognition, and the remaining two executive functioning tasks. Free recall began by establishing rapport with participants (e.g., asking participants about their day, about school, etc), after which the interviewer ensured that each participant understood the interview questions related to Bowtie Playtime by asking participants to describe the activity session leader's bowtie. After the interviewer was confident that the participant was comfortable and ready to begin, she asked the participant, "Can you please tell me everything you remember about Bowtie Playtime?" To encourage complete recall of the activity session, the interviewer asked three non-directive prompts when the participant paused in his/her recall (e.g., "Is there anything else you can tell me about Bowtie Playtime?"). A series of recognition questions followed free recall. For each activity session detail, one recognition question related to the experienced option (e.g., "Did you wear a purple cape during Bowtie Playtime?") and one recognition question related to a distracter option (e.g., "Did you wear a red cape during Bowtie Playtime?"). Participants were given the choice of answering "Yes," "No," or "Don't Know" to each of the 40 recognition questions. Participants were randomly assigned to one of two random orders of recognition questions. All interviews were audiotaped and, with permission, videotaped. Interviews were conducted by one of four trained interviewers. See Appendix B for an example of the second interview.

**Coding.** Interviewers were instructed to record participants' responses in detail to enable coding of each critical detail reported by participants. Accuracy of children's free recall and recognition responses were coded separately. Each critical detail recalled by a participant in free recall was coded as either correct or incorrect. A correct response occurred when a child described the experienced critical detail in specific or general terms (e.g., "purple apron" or "apron"), whereas an incorrect response was when a child described a critical detail that was not experienced during the activity session (e.g., "orange apron" or "cape"). Inter-coder agreement {[# of agreements / (# of agreements + # of disagreements)] × 100%} was 93.0% based on 15% of the sample.

Recognition responses were categorized as correct "Yes" responses (child responded "Yes" to questions about experienced critical details; e.g., "Did you wear a purple apron?"), incorrect "Yes" responses (child responded "Yes" to questions about non-experienced critical details; e.g., "Did you wear an orange apron?"), correct "No" responses (child responded "No" to questions about non-experienced critical details), incorrect "No" responses (child responded "No" to questions about experienced critical details), or "Don't Know" responses.

The consistency of recall attempts by children in the intervening interview condition was coded as well. Specifically, participants' descriptions of critical details in the first recall attempt were compared to their descriptions of critical details in the second recall attempt. There were six consistency coding categories:

1. Consistent: child's description of a critical detail did not change across recall attempts (e.g., "purple apron" and "purple apron")
2. Omission first recall: child described critical detail in the second recall only

3. Omission second recall: child described critical detail in the first recall only
4. Increased specificity: child provided more information about critical detail across recall attempts (e.g., "apron" and "purple apron")
5. Decreased specificity: child provided less information about critical detail across recall attempts (e.g., "purple apron" and "apron")
6. Contradictory: child's description of critical detail changed across recall attempts (e.g., "purple apron" and "orange apron")

Inter-coder agreement was 96.6% based on 15% of the sample.

# RESULTS

## *Directed Forgetting*

To examine whether DF instructions can affect reports of an autobiographical event, analyses of variance (ANOVA) were conducted on number of details accurately described by participants. Separate ANOVAs were conducted for List 1 and List 2 details to determine whether costs and benefits typically found in list method studies of DF were observed in the current study. Less than 10% of the sample recalled details not experienced during the activity session (i.e., incorrect responses); these responses were not analyzed further. All analyses were two-tailed.

**Interview 1.** See Table 2 for descriptive statistics for List 1 and List 2 recall during the first interview. A 4 (Age) × 2 (DF) ANOVA was conducted on number of List 1 details accurately recalled during the interview six days after the activity session. There were significant main effects of Age [$F(3, 113) = 17.26, p < .001, \eta^2 = .31$] and DF [$F(1, 113) = 4.98, p < .05, \eta^2 = .04$]. LSD post-hoc comparisons revealed that children in grades one and three reported the fewest number of List 1 details accurately, followed by children in grade five, and then university undergraduate students. No other comparisons were significant. With respect to DF, participants recalled more List 1 details when they were remember-cued than when they were forget-cued.

The Age × DF interaction was non-significant [$F(3, 113) = 0.43, p = .73, \eta^2 = .01$], suggesting the influence of DF instructions on recall during the first interview did not vary as a function of age. This was surprising, particularly considering empirical evidence that suggests list method DF effects often do not surface until about 10 years of age (e.g., Bray et al., 1983). Indeed, it was hypothesized that DF instructions would not influence younger children's recall of the activity session. Thus, a separate independent samples *t* test was conducted for each age group in spite of the non-significant interaction. These analyses were non-significant: Participants of all ages recalled as many List 1 details when they were remember-cued as when they were forget-cued (all *ps* > .10). Thus, although there was an overall effect of DF instructions, the effect was not significant within any age group. Rather, for each age group, the pattern of means was in the same direction, yielding an overall significant main effect of DF instructions (see Figure 1).[1]

A 4 (Age) × 2 (DF) ANOVA on number of List 2 details accurately recalled during the first interview revealed a significant main effect of Age only, $F(3, 113) = 12.49, p < .001, \eta^2 = .25$. LSD post-hoc comparisons revealed that children in grade one recalled fewer List 2 details accurately than children in grade five and adults. Children in grade three accurately reported as many List 2 details as children in grades one and five. Benefits of the forget instruction were not found, as the main effect of DF was non-significant, $F(1, 113) = 0.03, p = .86, \eta^2 = .01$.

**Interview 2.** See Table 3 for descriptive statistics for List 1 and List 2 recall

---

[1] That none of the independent samples *t* tests for age were significant was surprising as well. Given a priori expectations that DF would vary as a function of age, combined with what appeared to be mean differences depicted in Figure 1, planned contrasts were conducted. According to Howell (2002), these contrasts use the pooled variance estimate ($MS_{error}$) from the overall analysis of variance and evaluate *t* on the $df_{error}$ degrees of freedom. In other words, planned comparisons may be more sensitive and family-wise error rates are less problematic when only a few comparisons are made. These tests revealed that none of the children demonstrated significant DF effects (all *t*'s < 1.96, the $t_{critical}$ for $t(113), p < .05$). However, the difference between number of List 1 items recalled by adults in the FR condition and adults in the RR condition approached significance, $t(113) = 1.84$. This suggests the overall main effect of DF for recall of List 1 details may be driven by the adults, but the lack of significant interaction implies that the effect for adults was not significantly greater than that for the younger groups.

during the second interview. All analyses conducted for Interview 2 included comparisons of whether participants had recalled the activity session during the first interview, resulting in two separate 4 (Age) × 2 (DF) × 2 (Interview) ANOVAs for List 1 and List 2. There were significant main effects of Age [$F(3, 219) = 39.45, p < .001, \eta^2 = .35$], DF [$F(1, 219) = 7.73, p < .01, \eta^2 = .03$], and Interview [$F(1, 219) = 8.05, p < .01, \eta^2 = .04$] for number of List 1 details recalled accurately in the second interview. With respect to the main effect of Age, all LSD post-hoc comparisons were significant, such that children in grade one reported the fewest number of List 1 details, followed by children in grade three, children in grade five, and then adult participants. As in the first interview, costs of the DF instructions were observed: Participants provided fewer accurate descriptions of List 1 details when they were forget-cued than when they were remember-cued. With respect to the main effect of Interview, participants who recalled the activity session for the second time during the second interview recalled more List 1 details than participants who were recalling the activity session for the first time. The DF × Interview interaction was non-significant [$F(1, 219) = 0.11, p = .74, \eta^2 = .00$], suggesting that repeated recall attempts did not attenuate costs associated with the forget cue. In other words, participants who recalled the activity session in the first interview still evinced DF costs in the second interview. This suggests that repeated recall attempts do not provide a release from inhibition, which is examined further in a forthcoming section below.[2]

[2] The null DF × Interview interaction for Interview 2 was not surprising considering the planned comparisons for Interview 1 revealed that only adult responses were approaching significant DF costs. Thus, a 2 (DF) × 2 (Interview) ANOVA was conducted on adult responses only. The DF × Interview interaction for this analysis was non-significant as well, $F(1, 60) = 0.12, p = .73$. This provides further evidence that participation in an intervening interview did not lead to a reduction of DF effects in subsequent interviews.

The Age × DF interaction was non-significant in the second interview as well, $F(3, 219) = 1.59, p = .19, \eta^2 = .02$. Given the a priori hypothesis regarding developmental differences, follow-up independent samples $t$ tests were conducted similar to those conducted for responses given during the first interview. Findings replicated those from the developmental DF literature. Children in grades one and three did not demonstrate DF costs: Younger children were just as likely to recall List 1 details when they were remember-cued as when they were forget-cued [$t(56) = 0.26, p = .79$, $t(55) = 0.73, p = .47$ for grade one and grade three, respectively]. Conversely, children in grade five and adults recalled more List 1 details when they were remember-cued than when they were forget-cued [$t(54) = 2.12, p < .05$, $t(62) = 2.31, p < .05$ for grade five students and adults, respectively]. See Figure 2 for a depiction of means.

For recall of List 2 details, there was a significant main effect of Age only, $F(3, 219) = 25.86, p < .001, \eta^2 = .26$. LSD post-hoc comparisons revealed that children in grade one provided the fewest accurate descriptions of List 2 details, followed by children in grade three, and then children in grade five and adults. Although there was no significant main effect of DF [$F(1, 219) = 0.06, p = .81, \eta^2 = .00$], there was a significant Age × DF interaction, $F(3, 219) = 6.46, p < .05, \eta^2 = .04$. Separate independent samples $t$-tests for each age group revealed that adults in the RR condition recalled more List 2 details than adults in the FR condition [$t(62) = 2.26, p < .05$], which is opposite to what is expected for a DF benefit. No significant effects were found for child participants (all $p$`s $> .20$).

Recall that the second interview included 40 "Yes/No/Don't Know" recognition questions as well: one experienced option for each critical detail (e.g., "Did you wear a purple apron?") and one foil option for each critical detail (e.g., "Did you wear an orange apron?"). Thus, 4 (Age) × 2 (DF) × 2 (Interview) ANOVAs were conducted on number of correct "Yes" responses provided for List 1 and List 2 details, separately (see Table 4 for descriptive statistics). Only correct "Yes" responses, a child's affirmation that a detail was experienced during the activity session, were analyzed because these responses are most relevant to examining what children are able to recall about an experience. Incorrect "Yes" responses were not analyzed because the vast majority (72%) of the sample did not provide any "Don't Know" responses and, because all participants were asked the same 40 recognition questions, analysis of incorrect "Yes" responses would have elicited a pattern of findings that mirrored those for the correct "Yes" responses. For recognition responses to List 1 details, there was a significant main effect of Age, $F(3, 221) = 3.91, p = .01, \eta^2 = .05$. LSD post-hoc comparisons revealed that children in grade one provided fewer correct "Yes" responses to recognition questions than children in grade five and adults. Children in grade three provided a similar number of correct "Yes" responses as all other participants. No other comparisons were significant. As typically found in list method DF studies, there was no main effect of DF instructions on recognition responses, $F(1, 221) = 0.09, p = .77, \eta^2 = .00$.

Recognition responses to List 2 details followed a similar pattern as described above. There was a significant main effect of Age for correct "Yes" responses, $F(3, 221) = 6.25, p < .001, \eta^2 = .08$. LSD post-hoc comparisons revealed that children in grade one provided fewer correct responses than children in grade five and adults. In addition, children in grade three provided fewer correct responses than grade five students, but not undergraduate students. No other comparisons were significant. The DF instructions did not affect recognition responses to List 2 details, $F(1, 221) = 0.23, p = .63, \eta^2 = .00$.

### *Executive Functioning*

**Manipulation Checks.** Recall that participants completed one task each to measure shifting (i.e., Making Trails), updating (i.e., Keep Track), working memory (i.e., Digit Backwards) and inhibitory control (i.e., Stroop Color-Word). The shifting and inhibition tasks included a control trial and a manipulated trial. These tasks are effective only when participants' performance on manipulated trials is slower than their performance on control trials. Thus, a 4 (Age) × 2 (Trial: Trail A, Trail B) repeated measures ANOVA was conducted on the number of items connected per second on Trail A (i.e., connecting numbers 1-15 sequentially) and Trail B (i.e., connecting numbers 1-13 and letters A-L in alternating sequence) of the Making Trails task. As expected, there was a significant main effect of Trial, such that participants connected more circles per second on Trail A ($M = 0.86, SD = 0.46$) than on Trail B ($M = 0.27, SD = 0.17$), $F(1, 234) = 773.92, p < .001, \eta^2 = .77$. Interestingly, there was also a significant Age × Trial interaction, $F(3, 234) = 20.87, p < .001, \eta^2 = .21$. A series of paired samples $t$ tests revealed significantly slower performance on Trail B than Trail A for all age groups: grade one [$t(59) = 14.71, p < .001$], grade three [$t(56) = 16.82, p < .001$], grade five [$t(56) = 13.03, p < .001$], and adults [$t(63) = 15.05, p < .001$]. A one-way ANOVA conducted on the difference scores (number of items connected per second on Trail A – number of items connected per second on Trail B) indicated a significant main effect of Age, $F(3, 234) = 20.87, p < .001, \eta^2 = .21$. LSD post-hoc comparisons revealed that the difference between Trail A and Trail B performance for children in grade one was smaller than for all other participants. Difference scores for children in grades three and five did not differ, but both were smaller than the difference scores for adults. These findings suggest that adults completed the Making Trails task more efficiently than the child participants, the least efficient of which were children in grade one.

A 4 (Age) × 2 (Trial: Color, Color-Word) repeated measures ANOVA was conducted on the number of items identified per second on the Stroop Color-Word test. Only performance on the Color and Color-Word trials was examined for two reasons: (1) the measure of inhibitory control was calculated from these trials (inhibitory control = CW – C), and (2) previous research often administers these two trials only, as they best represent the control and manipulated tasks of the Stroop test (e.g., van der Sluis et al., 2007). As with the Making Trails task there was a significant main effect of Trial, such that participants identified more items per second on the Color task ($M = 1.23$, $SD = 0.39$) than on the Color-Word task ($M = 0.72$, $SD = 0.30$), $F(1, 231) = 1478.15$, $p < .001$, $\eta^2 = .87$. There was also a significant Age × Trial interaction, $F(3, 231) = 10.39$, $p < .001$, $\eta^2 = .12$. A series of paired samples *t* tests revealed significantly slower performance on the Color-Word trial than the Color trial for all age groups: grade one [$t(57) = 17.93$, $p < .001$], grade three [$t(56) = 18.80$, $p < .001$], grade five [$t(55) = 18.28$, $p < .001$], and adults [$t(63) = 22.27$, $p < .001$]. As with the Making Trails task, a one-way ANOVA conducted on the difference scores (number of items identified per second on the Color trial – number of items identified per second on the Color-Word trial) indicated a significant main effect of Age, $F(3, 231) = 10.30$, $p < .001$, $\eta^2 = .12$. LSD post-hoc comparisons revealed that the difference between the number of items identified in the Color and Color-Word trials was smaller for children in grade one than for children in grade five and adults. Children in grade three performed similarly to children in grades one and five, but their difference scores were smaller than those found for adults. Children in grade five and adults performed similarly.

By all of these indicators, the manipulated trials for the Making Trails and Stroop Color-Word tasks showed the expected decelerating effect on the participants' speed, suggesting that the shifting and inhibitory manipulations were effective. For all subsequent analyses involving EF, only manipulated trials for the Making Trails (i.e., Trail B) and inhibitory control scores are included (but see Table 5 for descriptive information on all trials of the EF tasks).

**Developmental Differences.** Separate one-way (Age) ANOVAs were conducted to examine developmental differences in updating (e.g., Keep Track), shifting (e.g., Making Trails), working memory (i.e., Digit Backwards), and inhibitory control (i.e., Stroop Effect). Significant main effects of Age were found for all EF tasks: Trail B [$F(3, 234) = 156.02, p < .001, \eta^2 = .67$], Keep Track [$F(3, 242) = 16.83, p < .001, \eta^2 = .17$], Digit Backwards [$F(3, 238) = 20.89, p < .001, \eta^2 = .21$], and the Stroop inhibitory control score [$F(3, 231) = 110.30, p < .001, \eta^2 = .12$]. Overall, and as expected, older participants performed better on the EF measures than the younger participants. All LSD post-hoc comparisons for Trail B were significant: Children in grade one connected the fewest items per second, followed by children in grade three, children in grade five, and then adult participants. Using proportion correct, performance on the Keep Track and Digit Backwards tasks followed a similar pattern: Children in grade one had the smallest proportion of correct responses than all other participants, and children in grades three and five (who did not differ from each other) performed more poorly than adults. LSD post-hoc comparisons for Stroop inhibitory control scores indicated that children in grade one had smaller scores than children in grade five and adults, and children in grade three had smaller scores than adults. No other comparisons were significant.

Tables 6-9 depict correlations between EF tasks separated by age group. Tasks were not correlated for any age groups, with one exception: There was a significant relationship between Trail B and Stroop inhibitory control scores for children in grade one, $r = -.28$. Children who took longer to complete Trail B exhibited lower inhibitory control. These correlations indicate that the EF tasks were unrelated, which suggests that each task measured an executive function distinct from that measured by the other tasks.

**Regression Analyses.** A standard multiple regression was performed between remember-forget difference scores as the dependent variable and performance on the EF tasks as predictors. Remember-forget difference scores (i.e., number of List 2 details recalled correctly – number of List 1 details recalled correctly) were calculated from details reported in each participant's first free recall attempt. For some participants this was the interview conducted six days after the activity session and for other participants this was the interview conducted seven days after the activity session.[3] EF task performance was measured in the following ways: number of items connected per second on Trail B, proportion correct responses on the Keep Track and Digit Backwards tasks, and the Stroop inhibitory control score. Regression analyses were conducted with participants in the FR condition only, as the relationship between EF and recall of an event after receiving an instruction to forget was of primary concern for the current study. The model was non-significant, accounting for only 2.0% of the variability in remember-

[3] An independent samples $t$ test revealed no significant difference between R-F difference scores for participants who first recalled the event six days later ($M = -0.15$, $SD = 1.74$) and R-F difference scores for participants who first recalled the event seven days later ($M = 0.22$, $SD = 1.61$), $t(124) = 1.22$, $p = .22$.

forget difference scores (see Tables 10 and 11 for regression coefficients and correlations, respectively).[4]

### *Repeated Interviews*

**Release from inhibition.** To examine whether repeated recall attempts provided a release from inhibition, 4 (Age) × 2 (DF) × 2 (Recall Attempt: first, second) repeated measures ANOVAs were conducted on number of details recalled correctly from List 1. Analyses were conducted on only those participants who recalled the activity session during both interviews and Recall Attempt served as a within-subjects variable. If repeated recall attempts released participants from retrieval inhibition, then a significant DF × Recall Attempt interaction was expected such that participants instructed to forget List 1 (i.e., FR condition) would demonstrate a larger increase in number of List 1 details reported across interviews than participants instructed to remember List 1 (i.e., RR condition). Thus, analysis of List 1 recall was required only. There were significant main effects of Age [$F(3, 105) = 20.41, p < .001, \eta^2 = .37$] and DF [$F(1, 105) = 5.31, p < .05, \eta^2 = .05$]. LSD post-hoc comparisons revealed that children in grades one and three provided the fewest accurate descriptions of List 1 details (*M*s = 5.28, 5.78, *SD*s = 2.36, 2.15, respectively), followed by children in grade five (*M* = 8.08, *SD* = 2.45) and adults (*M* = 10.16, *SD* = 3.58), who differed from each other. As in the analyses conducted with the entire sample, participants in the FR condition recalled fewer List 1 details (*M* = 6.77, *SD* = 3.11) than participants in the RR condition (*M* = 8.04, *SD* = 3.54). The DF × Recall interaction was non-significant [$F(1, 105) = 1.67, p = .20, \eta^2 = .02$], providing additional

[4] The regression analyses remained unchanged using proportion scores for the Making Trails task, the Stroop Color-Word test, and the R-F difference criterion value.

evidence that repeated recall attempts did not provide a "release from inhibition". In other words, increases in the number of List 1 items reported across interviews were equal for participants in the FR and RR conditions.[5]

**Report consistency.** To evaluate the consistency of participants' recall of the activity session, descriptions of critical details were compared across the two recall attempts to create a response pair for each critical detail described. Thus, only responses from those participants who recalled the activity session in both interviews are included in these analyses. Upon examination of the frequency with which participants' responses fell into each of the six consistency coding categories (i.e., consistent, omission first interview, omission second interview, increased specificity, decreased specificity, and contradictory), it became apparent that less than 10% of the sample provided contradictory response pairs. These responses were not analyzed further. Moreover, given that the focus of the present study was on the influence of DF instructions on consistency and omissions, only three types of response pairs were analyzed: consistent, omission first recall, and omission second recall. As can be seen from Tables 12 and 13, the vast majority of participants' response pairs fell into one of these three categories. Specifically, the proportion of response pairs that decreased or increased in specificity ranged from 0.0% to 11.0% and so very little data were lost in analyzing only consistencies and response pairs that included an omission in one recall attempt. Separate 4 (Age) × 2 (DF) ANOVAs were conducted on proportion consistent, proportion omission first recall, and proportion omission second recall response pairs. This was done

[5] Given that only adult participants demonstrated DF effects in both interviews, a separate 2 (DF) × 2 (Recall Attempt) repeated measures ANOVA was conducted on number of List 1 details recalled correctly by adults. The interaction was non-significant, $F(1, 30) = 0.03$, $p = .86$. This provides additional evidence to suggest that FR participants were not more likely to recall previously unreported List 1 details in the second interview than RR participants.

for List 1 and List 2 details separately. These analyses were conducted on proportions to account for the fact that participants provided a different number of response pairs across recall attempts. All analyses were two-tailed.

For List 1 details, there was a significant main effect of Age for proportion consistent responses, $F(3, 106) = 7.02, p < .001, \eta^2 = .17$. LSD post-hoc comparisons revealed that children in grade one provided a smaller proportion of consistent responses than children in grade three, children in grade five, and adults. No other comparisons were significant. There were no significant main effects for omission first recall or omission second recall response pairs (all *p's* > .10).

There were no significant effects for List 2 details (all *p's* > .05).

# DISCUSSION

The present research was an innovative application of the directed forgetting (DF) procedure to memory for a personally experienced event. Specifically, list method DF instructions were issued twice during an interactive activity session: once halfway through the session and again at the end of the session. Participants' memory for the activities was evaluated via recall and recognition questions. Executive functioning and the consistency of participants' recall across two recall attempts were also measured. There were three primary hypotheses: (1) DF instructions would influence reports of a personally experienced event for older participants only (i.e., older children and adults), (2) participants with higher scores on shifting, updating, working memory, and inhibitory control would show larger DF effects, and (3) repeated recall attempts would reduce the effects of DF instructions on autobiographical memory by increasing recall of forget-cued details across interviews. Each of these hypotheses, and corresponding results, is discussed in the sections that follow.

## *Directed Forgetting*

**Theoretical explanations.** Findings in the current study replicated those reported in the two published applications of DF to memory for autobiographical events; specifically, costs with no benefits (Barnier et al., 2007; Joslyn & Oakes, 2005). Costs of the forget cue were observed such that participants recalled fewer List 1 activity details when they were forget-cued than when they were remember-cued. Follow-up analyses conducted on responses provided during the first interview suggested that responses from participants of all ages were in the direction of DF costs, but only adult responses approached significance. Analyses conducted on recall performance during the second interview showed DF costs for grade five students and adults only. There was no evidence of recall benefits in either interview. Overall, these findings demonstrate that list method DF instructions can extend to memory for a complex, interactive event in ways similar to memory for words in a list, sentence, or story. These promising data may provide support for arguments that DF can be used to examine forgetting of autobiographical memories (e.g., Anderson & Green, 2001; Koustaal & Schacter, 1997). Recall that there are currently three possible explanations for list method DF effects: inhibition, selective rehearsal, and a two-factor model proposing separate mechanisms for costs (context change) and benefits (strategy switch). Each of these explanations relies on an interpretation of both recall and recognition responses, as well as cost and benefit analyses. Thus, to explore possible underlying mechanisms responsible for the current study's findings, what follows is a description of each explanation and relevant data.

Until recently, many scholars in the DF literature argued that inhibition accounted for DF effects using the list method of cue presentation. According to this explanation, an instruction to forget initiates a process that inhibits access to List 1 items during recall (e.g., E. L. Bjork & R. A. Bjork, 1996; R. A. Bjork, 1989). One of the most persuasive pieces of evidence used to support the inhibition account of list method DF findings is that DF effects typically disappear in response to recognition questions, when item generation is not required. This evidence is made more compelling by the fact that many scholars who argue against inhibition focus on recall performance only (e.g., MacLeod et al., 2003; Sahakyan et al., 2004). Findings garnered from the current study were not unlike those found in typical list method studies: DF instructions led to costs in free recall reports of the activity session but did not appear to influence participants' recognition responses. Thus, on their surface, the results of the current study appear to support the oft-cited inhibition account of list method DF effects.

Nevertheless, that there were no significant DF effects in recognition may not necessarily support inhibition operating in recall either. It is the case that item method studies ordinarily find a decrease in the size of the DF effect from recall to recognition, and that DF effects found in list method studies are often smaller than those found in item method studies (MacLeod et al., 2003). Taken together, if a similar reduction in the size of the DF effect occurs from recall to recognition in list method studies as in item method studies, then the effect size in recognition for list method studies may be at the floor and so unobservable. To know if recognition merely reduces the size of the DF effect or eliminates it, it is necessary to observe a medium to large DF effect in recall. This was not the case in the present study, as the observed effect sizes for the main effects of DF found in recall were quite small ($\eta^2$ = .04, .03 for Interview 1 and 2, respectively). Furthermore, previous studies have demonstrated that forgetting effects may be smaller for autobiographical memories than memory for word lists (Barnier et al., 2007), making a floor effect in recognition more likely. Thus, these data cannot rule out alternative explanations for list method DF effects, particularly selective rehearsal.

MacLeod and colleagues (2003) presented a series of studies that highlighted the potential role of selective rehearsal in list method findings. Their findings indicated that increasing the retention interval between item presentation and recall led to increased DF effects compared to a standard DF condition without a delay to test. Interestingly, when participants were warned that the upcoming recall task required recall of forget-cued details as well as remember-cued details, the resulting DF effects were smaller than those observed in the standard DF condition without a delay to test. MacLeod and colleagues argued that selective rehearsal explained this pattern of findings. They believed that, during the delay, participants not warned of the upcoming recall test focused their rehearsal efforts on remember-cued details. They had no reason to expect a request to report forget-cued details, which led to larger DF effects. Conversely, MacLeod and colleagues proposed that participants warned about the content of the recall task focused their rehearsal efforts on forget-cued details relative to remember-cued details, which led to smaller DF effects. Participants in the current study were asked to recall the activity session six and/or seven days later, providing them with the opportunity to selectively rehearse remember-cued items more than forget-cued items. Thus, it is reasonable to suggest that DF costs found in the current study may be the result of selective rehearsal as opposed to inhibition. However, this is merely speculative. The current study did not include a retention interval manipulation and therefore the size of the DF effect at different delays could not be compared directly. An extension of those studies described by MacLeod and colleagues using autobiographical events and longer retention intervals may be an interesting direction for future research.

Recall that Sahakyan and colleagues (e.g., 2003; 2005) proposed a two-factor model for list method costs (context change) and benefits (strategy shift). Findings from the current study may point to a two-factor model in that costs were observed but not benefits. According to the *context change hypothesis* (Sahakyan & Kelley, 2002), costs arise because participants encode List 2 items within a new mental context that better matches the retrieval context. This argument implies that the encoding context of List 1 is different from that of List 2. This may have been possible in the current study because List 1 activities were distinct from List 2 activities; however, additional studies that specifically encourage context change are necessary to test fully the application of the *context change hypothesis* to DF effects in autobiographical memory reports.

According to the strategy-shift proposition (Sahakyan & Delaney, 2005), DF benefits were not observed in the current study because participants failed to change their encoding strategy for List 2 activities. This may be true, in part, because of the way the study was explained to participants. At the beginning of each activity session, the experimenter introduced the study as an investigation of "how people remember events". This may have encouraged both the FR and RR groups to utilize effective study strategies from the outset, reducing the likelihood that participants in the FR condition switched to more effective study strategies after hearing the forget instruction. Nonetheless, these findings are not unlike those described in previous studies of DF and autobiographical memory (Barnier et al., 2007; Joslyn & Oakes, 2005), and may indicate that benefits are less likely to occur for recall of autobiographical memories. Specifically, characteristics of the activity session may not have encouraged the use of study strategies in the same way that learning a list of words might. First, activity sessions were conducted with groups of participants so that activities were more engaging and interactive. Second, the play session leader facilitated encoding of the critical details by verbally naming each detail approximately three times. Third, in most cases there were multiple copies of each item so that participants could interact with the details. Thus, overt study strategies may not be required to encode an interactive activity session compared to when participants study words presented on a computer screen. Overall, the methods employed in the current study, make application of the current findings to a two-factor model of DF effects difficult.

In summary, findings from the current study were similar to those typically found in DF studies using the list method: DF effects were observed in recall but not recognition. Historically, this pattern of findings has been used to advance the proposition that retrieval inhibition is responsible for list method findings. Although data from the current study does not provide unequivocal evidence against retrieval inhibition, selective rehearsal cannot be ruled out entirely. It is reasonable to believe that the experimenter's explanation of the study to participants and the retention interval between the activity session and memory interview(s) may have facilitated selective rehearsal. The two-factor account for DF effects may explain the pattern of findings in the current study as well, namely that costs were observed without corresponding benefits. The initial motivation for the current study was not to test underlying mechanisms and provide definitive evidence in favour of one mechanism over another. Rather, the current study intended to create a procedure that led to DF effects in reports of a personally experienced event that could be used in future studies. One interesting extension of this methodology could include a systematic evaluation of underlying mechanisms.

**Developmental differences.** Interestingly, initial analyses did not uncover developmental differences in DF in either interview. This was unexpected given a priori hypotheses for the current study based on previous literature on the development of DF (see Wilson & Kipp, 1998 for a review). Follow-up analyses conducted on recall provided during the first interview revealed a near-significant DF effect for adult participants, but non-significant DF effects for all child participants. This may have occurred because only half of the participants recalled the activity session during the first interview. Given the reported small effect of DF instructions on recall, there may have been insufficient power to detect developmental differences. Indeed, the observed power for the Age × DF interaction for responses provided in the first interview was .13. In the second interview recall, however, follow-up analyses revealed the expected age differences. Whereas older participants (i.e., children in grade five and adults) demonstrated typical list method costs, younger participants (i.e., children in grades one and three) were unaffected by the forget instruction and recalled as many List 1 details when they were forget-cued as when they were remember-cued. These findings are consistent with previous developmental studies in which children do not evince list method DF effects until grade five (e.g., Bray et al., 1993). Typically, forget-cued items appear in young children's recall more often than in older children's and adult's recall. Scholars have proposed these developmental differences may, in part, be due to children's developing selective rehearsal and inhibition skills (e.g., Lehman et al., 2001). One of the primary objectives for assessing participants' executive functioning in the current study was to explore whether executive functions contribute to the development of DF, which is discussed in more detail in the upcoming section.

A common application of DF to autobiographical memory, and motivation for the current study, is to memory for childhood sexual abuse (CSA). Several scholars have put forth the notion that avoiding thoughts about the abuse may lead to DF effects when reporting the abuse (e.g., Anderson & Green, 2001; Epstein & Bottoms, 2002). However, their proposal is based upon a body of research using primarily adult participants. This study was the first empirical investigation into the influence of DF on *children's* memory for a personally experienced event. Findings in the current study suggest that DF does not influence younger children's reports of an interactive activity session, which is consistent with some studies of descriptive reports of childhood sexual abuse. In a survey of 129 women with documented histories of CSA, Williams (1995) found that some women did not immediately begin "blocking out" thoughts of their abuse experiences when they were children. Rather, women reported active avoidance later on in life. Unfortunately, whether older children were more likely to avoid thoughts of the abuse than younger children could not be determined from the data presented by Williams. Social factors related to children's subjective experiences of the abuse may reduce the likelihood that children will actively avoid thoughts of the abuse as well. For example, children may not initially perceive abuse as a negative and socially reprehensible act. Sas and Cunningham (1995) found that 40% of 135 children interviewed after participating in court proceedings reported "not knowing the abuse was wrong" as a factor contributing to their delayed disclosure. In addition, the younger the children were at the time the abuse occurred, the more likely they were to report this as their reason for non-disclosure. It is not proposed that findings from the current study provide definitive evidence against the possible application of DF to childhood autobiographical memories. Indeed, older children (grade five students) were able to direct their forgetting of the activity session. Instead, for reasons described above, there may be limitations of such an application that warrant additional investigation.

## *Executive Functioning*

The role of individual differences in DF effects is virtually unexplored and the inclusion of executive functioning (EF) measures in the current study afforded the opportunity to investigate whether developmental differences in DF are related to other developing cognitive skills. Contrary to hypotheses, EF task performance was not related to the effect of DF instructions on recall. Using the remember-forget difference score as the criterion variable, regression analyses showed that none of the EF measures predicted FR participants' responses to the forget cue. It is possible to conclude from these findings that EF does not relate to the ability to "forget" seemingly irrelevant information (i.e., forget-cued information). However, before that can be presented as a plausible explanation, it is important to consider other reasons for failure to observe a relationship.

The problem of "task impurity" (e.g., Denckla, 1994; Rabbitt, 1997) may help to explain why EF was not related to DF in the current study. As discussed before, measurement of EF is notoriously difficult, made especially so by unreliable and "impure" tasks. Many tasks that purport to measure a particular EF may in fact require other cognitive skills (e.g., verbal ability, motor speed) or other executive functions for successful completion. For this reason, several steps were undertaken to ensure proper measurement of EF. First, the executive functions measured in the current study were considered carefully. Shifting, updating, and inhibition were chosen because they have been identified as key executive functions in the literature (e.g., Miyake et al., 2000), are well-defined relative to other executive functions (e.g., Baddeley, 1996; Rabbitt; van der Sluis et al., 2007), and seemed most relevant to skills implicated in DF (e.g., inhibition). Many EF tasks call on working memory for successful performance, some even suggest it is an executive function, and so a measure of working memory was included as well. Second, numerous tasks were researched in order to find the most appropriate measures of each executive function and working memory. It was imperative that tasks be suitable for child and adult participants, as well as demonstrate reliability in measurement and straightforward administration. All tasks included in the current study have been used with individuals as young as six years of age and older, are commonly used in the EF literature, and can be administered efficiently (e.g., Arffa, 2007; Brocki & Bohlin, 2004; Im-Bolter et al., 2006; St Clair-Thompson & Gathercole, 2006; van der Sluis et al.). Third, analyses were conducted to evaluate the effectiveness of the manipulations. Initial screening of the data suggested that measures of shifting and inhibitory control were successful, in that participants completed the manipulated trials significantly slower than the control trials. Fourth, as expected from the developmental psychology literature (e.g., Carlson, 2005; Mäntylä, Carelli, & Forman, 2007), performance on all EF tasks improved with age. It seems reasonable to conclude that the executive functions and corresponding tasks included in the current study were appropriate measures of shifting, updating, working memory, and inhibition.

In spite of the implemented safeguards, it is possible that task impurity affected proper measurement of EF in the current study. An oft-cited solution to task impurity is to include multiple tasks to measure a single executive function (e.g., van der Sluis et al., 2007). Originally, two tasks were included to measure each executive function and working memory. Unfortunately, during the initial stages of data collection, it became apparent that the youngest participants had difficulty completing the tasks in a timely manner. Participation from schools was solicited under the agreement that students would be outside of class for two 20-minute interviews, but some children took 45 minutes to complete each interview. Accordingly, half of the EF measures were excluded. Failure to demonstrate a relationship between EF and DF may be indicative of the task impurity problem, and unstable measurement of EF and working memory, rather than an actual lack of relationship. Future studies would benefit from including a more comprehensive battery of EF measures.

In addition to the aforementioned explanations for the lack of relationship between EF and DF, it is important to consider the possibility that the two constructs may in fact be unrelated. This explanation seems untenable based on the current findings. Specifically, the observed developmental differences were quite small and surfaced only after conducting follow-up tests in spite of non-significant interactions between Age and DF. Reduced variability in DF effects across age groups may have hindered the ability to detect a relationship with EF. This could be remedied in a number of different ways (e.g., measure age continuously, increase variability in amount of information recalled by using a shorter retention interval, by creating bigger lists, or by making items more salient). Thus, before discounting the possibility that EF contributes to DF effects, future research should attempt to enhance variability in DF across age groups.

### *Repeated Interviews*

The most common interpretation of list method DF findings is that inhibition accounts for DF effects in recall performance and, given that DF effects disappear when participants answer recognition questions, recognition provides a release from inhibition. However, if inhibition is responsible for DF effects, recognition is not a viable option for forensic interviewers. Rarely would an investigator possess knowledge of the crime necessary for creating a recognition test for an eyewitness and, even if that information was available, recognition questions may be considered leading, thereby tarnishing the quality of the evidence. A substantial body of literature has documented that repeated open-ended questioning can provide an inoculation against forgetting and lead to reminiscence and hypermnesia (e.g., Salmon & Pipe, 1997), suggesting that multiple recall attempts may facilitate a release from inhibition. It was expected that, if retrieval inhibition is responsible for DF findings using the list method, then participants might demonstrate increased recall of forget-cued details relative to remember-cued details across recall attempts. Thus, half of the participants in the current study were asked to recall the activity session twice: once six days after the activity session and again one day later. There was no compelling evidence that repeated recall attempts provided a release from inhibition. Participants who recalled the activity session twice recalled as many previously unreported forget-cued details in the second interview as previously unreported remember-cued details.

That repeated recall attempts did not lead to increased recall of forget-cued details in the second interview relative to recall of remember-cued details may lend support to a selective rehearsal interpretation of the findings. Recall that MacLeod and colleagues (2003) offered compelling evidence that selective rehearsal may play a role in list method DF effects over a delay of minutes. In their studies, those participants not warned about the upcoming recall test demonstrated greater DF effects after a delay than participants in the standard DF procedure condition. They argued that participants in the delay-no warning condition filled the delay rehearsing remember-cued items. Participants in the current study recalled the event six and seven days after the activity session (akin to a delay-no warning condition), which afforded participants in the FR condition plenty of time to focus their rehearsal efforts on the remember-cued details. If participants differentially rehearsed forget-cued and remember-cued details over the delay then we would not expect an additional recall attempt to facilitate their recall of the forget-cued details. Indeed, that was the case for the current study (see Goernert, 2005 for a similar unsuccessful attempt to provide a release from inhibition with repeated recall interviews).

What cannot be ignored is the difficulty associated with using null findings to support any claim. In other words, it is possible that methods used in the current study prevented repeated recall attempts from undoing the effects of directed forgetting. Indeed, the majority of children's response pairs were consistent, which may suggest that participants merely repeated information they reported in the first interview during the second interview one day later. That two different interviewers solicited recall from the participants may have encouraged this behaviour. Participants may not have attempted to recall more information during the second interview because they could just as easily report the same information from the first interview without the interviewer's knowledge. Thus, before concluding that repeated recall attempts cannot provide a release from inhibition, research addressing these methodological issues is necessary. For example, a future study could have the same interviewer conduct both interviews and actively encourage additional recall by providing incentives, praise, or reviewing what the participant recalled previously.

### *Limitations and Future Directions*

The current study was, of course, not without its limitations. One potential criticism of this research relates to external validity; specifically, whether findings from this study generalize to victims' memories for traumatic events. By framing justification for this research in the context of childhood sexual abuse (CSA), it is not presumed that DF is solely responsible for delayed disclosure of CSA, or that DF interferes with all victims' recollections of abuse. Moreover, the current work does not propose that DF leads to complete repression of or amnesia for abuse memories; in fact, only limited empirical evidence exists for so-called hidden memories of CSA (e.g., Goodman et al., 2003; Lindsay & Read, 1995; Williams, 1994). Rather, the overarching goal of the current study was to move toward an understanding of the effects of a long delay to disclosure by examining the influence of DF instructions on reports of an interactive play session.

There is ongoing debate regarding differences between memory for emotionally arousing events and memory for innocuous events, and whether the extensive literature on the latter can inform our understanding of the former (for recent reviews Deffenbacher, Bornstein, Penrod, & McGorty, 2004; Price & Connolly, 2008). In other words, although the intuitive link between DF and avoidance of upsetting memories is appealing, there may be limitations to the application of a forget instruction to motivated withholding of traumatic experiences like CSA. Arguably, victims of abuse possess clear motivation for non-disclosure (e.g., shame, embarrassment, fear of the criminal justice system, the abuser may share a close relationship with the victim) that is not present when participants recall an innocuous activity session for a friendly interviewer in a supportive environment. Nevertheless, recent research documents few differences between memory for emotionally arousing events and memory for neutral events. Price and Connolly (2007) developed a unique, and more experimentally controlled, way to compare children's memory for emotionally arousing events with their memory for neutral events. One problem associated with studying memory for emotionally arousing events relates to the absence of an appropriate comparison group; studies typically compare memory for two different events, making it impossible to decipher the relative influences of arousal and event type on recall. Price and Connolly addressed this confound by studying children's memory for swimming lessons, an experience that was stressful for some children (i.e., half of the children were fearful of water) and not stressful for other children. They found that children's accuracy when recalling the swimming lessons did not vary as a function of whether the swimming lessons were emotionally arousing. These findings highlight the utility of research on memory for neutral events, and that this work may indeed generalize beyond the laboratory. That said future research could explore the influence of DF instructions on memory for emotionally arousing events.

Time constraints prevented the inclusion of multiple tasks for shifting, updating, working memory, and inhibitory control. Thus, task impurity may have hindered proper measurement of EF. A study that focuses solely on the relationship between EF and DF, and allows longer interviewing times, would allow for inclusion of several EF tasks. In addition, it would be helpful to measure age continuously in lieu of including children from particular cohorts only. This would maximize variability in EF performance and DF effects, thereby increasing sensitivity in measuring these constructs.

Future research may also continue to explore the forensic implications of DF and apply the procedure to other issues relevant to child witnesses. An example of this includes memory for repeated events. Unfortunately, some studies report that at least half of CSA victims allege repeated abuse (Sas, 1993), but most autobiographical memory studies examine memory for unique events. The current study is no exception. There is a substantial body of literature to demonstrate that repeated events are represented in memory differently than unique events (eg., Hudson, 1990; Nelson & Gruendel, 1986). Research highlights these differences in several ways: children describing repeated events use more general language (e.g., Fivush, 1997), have difficulty attributing variable details to an instance (e.g., Powell & Roberts, 2002), are more suggestible (e.g., Connolly & Price, 2006; Price & Connolly, 2004), but are more accurate when describing details that do not vary across instances of the event (e.g., Connolly & Lindsay, 2001). At least two theories help to explain these differences. On the one hand, script theory argues that, with repeated exposure, a cognitive representation is formed for what usually happens during a routine (Nelson, 1986). Details that are fixed, or invariant, across repetitions are remembered well because they are represented in the script. Variable details are more difficult to remember because they are a dynamic list-like set of options that are not strongly associated with any one instance. On the other hand, fuzzy-trace theory (e.g., Brainerd & Reyna, 2002; 2004; Reyna, Holliday, & Marche, 2002) asserts that two independent memory traces are formed each time an event is experienced: a verbatim trace (specific information about each instance) and a gist trace (general meaning activated with each instance). Verbatim trace memory is stronger after repeated exposure to a detail (fixed details) than for details encountered only once (variable details). An interesting extension of this research would include issuing DF instructions during an instance of a repeated event comprised of both fixed and variable details. Results could inform the ongoing debate regarding underlying mechanisms responsible for the DF effect. If selective rehearsal accounts for DF effects then DF instructions should have little influence on recall of fixed details, as participants rehearse these details during every instance of the repeated event. An inhibition account would not predict that DF instructions would have a differential impact on fixed and variable details, as inhibition is not selective and would affect recall of the entire list of items.

# CONCLUSION

The current research is an important, and critical, first step in understanding the application of DF to children's autobiographical memory. The initial obstacle in beginning this line of research was creating a procedure that led to significant DF effects. Results from this study successfully replicated DF effects found in studies using word lists: Older participants provided fewer details about the first two activities when they were forget-cued than when they were remember-cued. In other words, older children and adults demonstrated difficulty recalling portions of an event after receiving an instruction to forget. These findings indicate that DF instructions can influence reports of personally experienced events and may generalize to arguments in favour of applying DF to examine forgetting of traumatic experiences like childhood sexual abuse, when motivation to direct forgetting is arguably much greater than in laboratory settings (e.g., Anderson & Green, 2001; Epstein & Bottoms, 2002; Koustaal & Schacter, 1997). Developmental differences found in the current study suggest that the youngest participants were unable to direct forgetting of the activity session, which may indicate that DF effects do not contribute to delayed disclosure for young victims. Instead, other factors may be more likely to play a role, such as failure to understand the inappropriateness of abuse. However, for those victims who may have difficulty retrieving memories for their abuse experiences because of some DF mechanism, it is imperative that proper interviewing techniques are available for forensic interviewers. These interviewing techniques should aim to solicit the most comprehensive and accurate report possible from a witness. The current study included a repeated interview condition as a way of exploring the possible benefit of repeated recall attempts on memory reports and did not find that recalling an event multiple times leads to a reduction in omissions. Future studies should examine conditions that maximize the benefits of repeated interviewing because repeated interviewing is commonplace in the forensic context and may provide, under appropriate circumstances, a way of reducing omissions in children's reports. By all accounts, the procedure created in the present study marks the beginning of an important area of research into the DF phenomenon and tests of the limitations of its application to autobiographical memory.

**Figure 1.** ***Mean (Standard Error Bars) Recall of List 1 Details as a Function of Age and DF Condition for Interview 1***

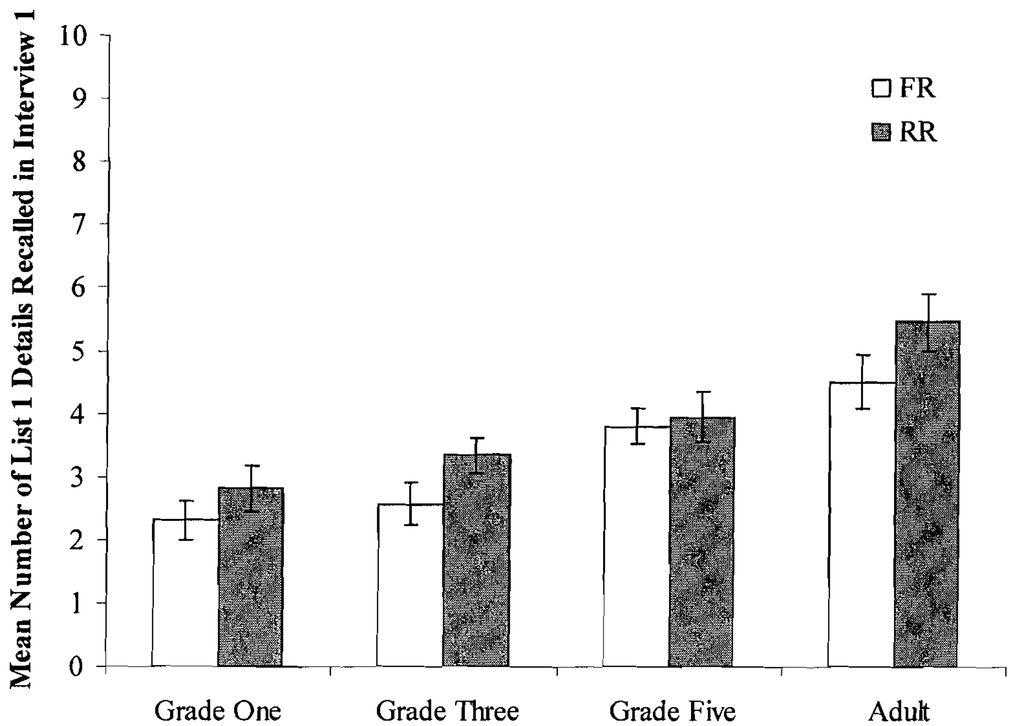

**Figure 2.** ***Mean (Standard Error Bars) Recall of List 1 Details as a Function of Age and DF Condition for Interview 2***

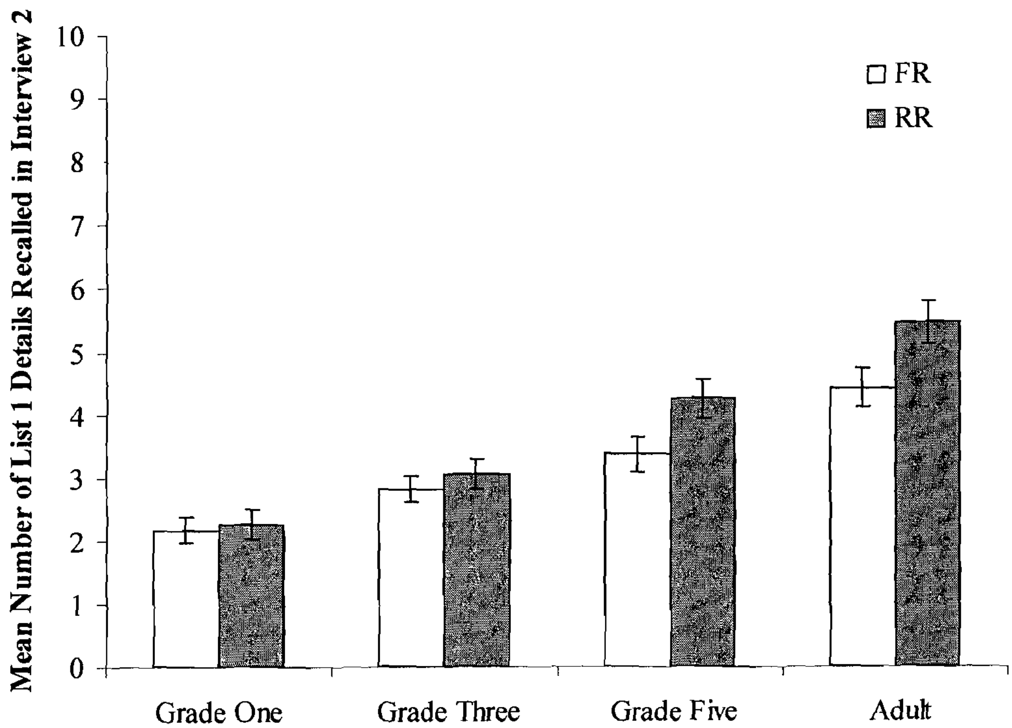

**Table 1.** ***List of Critical Details***

| Activity | Critical Details |
|---|---|
| *Paper Folding* | 1. *Kite* |
| | 2. Alex nametag |
| | 3. Think about favourite type of pizza |
| | 4. Lucky 4-leaf clover |
| | 5. Listen to piano music |
| *Magic Trick* | 1. *Disappearing Ball* |
| | 2. Straw hat |
| | 3. Sing "Row Row Row Your Boat" |
| | 4. Big Bird helper |
| | 5. Birthday party banner |
| *Sticker Colouring* | 1. *Airplane* |
| | 2. Ring |
| | 3. Jumping jacks exercise |
| | 4. Drink apple juice |
| | 5. $500 reward |
| *Clay Molding* | 1. *Flowers* |
| | 2. Purple apron |
| | 3. Say "Silly String" code words |
| | 4. Seashells |
| | 5. 10 o'clock in the morning |

**Table 2.** ***Means (SDs) of accurate List 1 and List 2 recall in Interview 1***

| | | List 1 (/10) | List 2 (/10) |
|---|---|---|---|
| Grade 1 | F-R | 2.31 (1.25) | 2.31 (1.20) |
| | R-R | 2.81 (1.47) | 2.06 (1.12) |
| | Overall | 2.56 (1.37) | 2.19 (1.15) |
| Grade 3 | F-R | 2.56 (1.37) | 2.50 (1.51) |
| | R-R | 3.33 (0.99) | 2.92 (1.24) |
| | Overall | 2.89 (1.26) | 2.68 (1.39) |
| Grade 5 | F-R | 3.79 (1.05) | 3.50 (1.70) |
| | R-R | 3.93 (1.53) | 3.33 (1.29) |
| | Overall | 3.86 (1.30) | 3.41 (1.48) |
| Adult | F-R | 4.50 (1.67) | 4.25 (1.24) |
| | R-R | 5.44 (1.86) | 4.44 (2.22) |
| | Overall | 4.97 (1.81) | 4.34 (1.77) |
| Overall | F-R | 3.27 (1.61) | 3.13 (1.59) |
| | R-R | 3.92 (1.80) | 3.20 (1.76) |

**Table 3.** ***Means (SDs) of accurate List 1 and List 2 recall in Interview 2***

| | | | List 1 (/10) | List 2 (/10) |
|---|---|---|---|---|
| Grade 1 | F-R | Single Interview | 1.75 (1.07) | 1.69 (1.20) |
| | F-R | Repeat Interview | 2.64 (1.08) | 2.57 (1.60) |
| | R-R | Single Interview | 1.69 (1.03) | 2.08 (1.04) |
| | R-R | Repeat Interview | 2.73 (1.28) | 2.13 (0.99) |
| | Overall | | 2.21 (1.20) | 2.10 (1.24) |
| Grade 3 | F-R | Single Interview | 2.88 (1.09) | 3.00 (1.83) |
| | F-R | Repeat Interview | 2.73 (1.28) | 2.93 (1.22) |
| | R-R | Single Interview | 2.86 (1.46) | 2.29 (1.38) |
| | R-R | Repeat Interview | 3.25 (0.87) | 3.25 (1.36) |
| | Overall | | 2.91 (1.18) | 2.86 (1.48) |
| Grade 5 | F-R | Single Interview | 2.94 (1.48) | 4.25 (1.29) |
| | F-R | Repeat Interview | 3.92 (1.44) | 4.33 (1.61) |
| | R-R | Single Interview | 4.20 (1.86) | 3.87 (2.13) |
| | R-R | Repeat Interview | 4.31 (1.38) | 3.62 (1.33) |
| | Overall | | 3.80 (1.62) | 4.02 (1.61) |

| | | | | |
|---|---|---|---|---|
| Adult | F-R | Single Interview | 4.06 (1.81) | 3.56 (1.37) |
| | F-R | Repeat Interview | 4.75 (1.61) | 4.06 (1.73) |
| | R-R | Single Interview | 5.25 (1.57) | 4.94 (1.57) |
| | R-R | Repeat Interview | 5.63 (2.13) | 4.63 (2.16) |
| | Overall | | 4.92 (1.85) | 4.30 (1.77) |
| Overall | F-R | | 3.20 (1.63) | 3.28 (1.69) |
| | R-R | | 3.82 (1.95) | 3.40 (1.87) |
| | Single Interview | | 3.24 (1.83) | 3.25 (1.82) |
| | Repeat Interview | | 3.78 (1.77) | 3.44 (1.72) |

**Table 4.** ***Means (SDs) of correct "Yes" Responses in Recognition***

| | | | List 1 (/10) | List 2 (/10) |
|---|---|---|---|---|
| Grade 1 | F-R | Single Interview | 7.88 (1.93) | 8.00 (1.46) |
| | F-R | Repeat Interview | 8.36 (0.75) | 8.14 (1.51) |
| | R-R | Single Interview | 7.13 (2.17) | 7.73 (1.87) |
| | R-R | Repeat Interview | 7.80 (1.52) | 8.07 (1.58) |
| | Overall | | 7.78 (1.71) | 7.98 (1.58) |
| Grade 3 | F-R | Single Interview | 8.75 (0.93) | 8.44 (1.32) |
| | F-R | Repeat Interview | 7.47 (1.55) | 8.27 (1.39) |
| | R-R | Single Interview | 8.29 (1.27) | 8.14 (1.46) |
| | R-R | Repeat Interview | 8.33 (0.99) | 8.00 (1.13) |
| | Overall | | 8.21 (1.28) | 8.23 (1.31) |
| Grade 5 | F-R | Single Interview | 8.31 (1.45) | 8.75 (0.93) |
| | F-R | Repeat Interview | 8.42 (1.24) | 9.25 (0.62) |
| | R-R | Single Interview | 9.00 (0.85) | 8.87 (0.92) |
| | R-R | Repeat Interview | 9.08 (0.86) | 8.85 (0.90) |
| | Overall | | 8.70 (1.15) | 8.91 (0.86) |

| | | | | |
|---|---|---|---|---|
| Adult | F-R | Single Interview | 8.50 (1.59) | 8.56 (1.15) |
| | F-R | Repeat Interview | 8.19 (1.28) | 8.31 (0.95) |
| | R-R | Single Interview | 8.38 (1.89) | 8.94 (1.00) |
| | R-R | Repeat Interview | 8.31 (1.62) | 8.50 (1.21) |
| | Overall | | 8.34 (1.58) | 8.58 (1.08) |
| Overall | F-R | | 8.23 (1.41) | 8.45 (1.22) |
| | R-R | | 8.28 (1.56) | 8.40 (1.34) |
| | Single Interview | | 8.28 (1.62) | 8.44 (1.32) |
| | Repeat Interview | | 8.22 (1.32) | 8.41 (1.24) |

**Table 5.** ***Descriptive statistics for Trail Making, Keep Track, Digit Backwards, and Stroop task performance***

| | Grade 1 | | Grade 3 | | Grade 5 | | Adult | |
|---|---|---|---|---|---|---|---|---|
| *Task* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Trail A | 0.50 | 0.03 | 0.74 | 0.03 | 0.86 | 0.04 | 1.32 | 0.06 |
| Trail B | 0.10 | 0.01 | 0.21 | 0.01 | 0.27 | 0.01 | 0.47 | 0.02 |
| Keep Track | 0.58 | 0.02 | 0.67 | 0.01 | 0.66 | 0.02 | 0.74 | 0.02 |
| Digit Backwards | 0.60 | 0.02 | 0.69 | 0.02 | 0.72 | 0.01 | 0.77 | 0.01 |
| Stroop Word | 1.18 | 0.04 | 1.62 | 0.03 | 1.79 | 0.03 | 2.24 | 0.03 |
| Stroop Color | 0.87 | 0.03 | 1.08 | 0.03 | 1.24 | 0.03 | 1.67 | 0.04 |
| Stroop Color-Word | 0.46 | 0.02 | 0.61 | 0.02 | 0.69 | 0.02 | 1.07 | 0.03 |
| Inhibitory Control | -18.47 | 7.85 | -21.49 | 8.63 | -24.57 | 10.06 | -23.03 | 9.66 |

*Note:* Scores on Trail Making and Stroop tasks are number of items connected or named per second. Scores on Keep Track and Digit Backwards tasks are proportion correct. The Inhibitory Control score is the value obtained after subtracting a participant's raw score on the color trial from their raw score on the color-word trial (i.e., SE = CW – C).

**Table 6.** ***Correlations between Trail Making (TB), Keep Track (KT), Digit Backwards (DB), and Stroop (S) tasks for Grade One students***

| | TB | KT | DB | S |
|---|---|---|---|---|
| TB | -- | | | |
| KT | .13 | -- | | |
| DB | .07 | -.03 | -- | |
| S | -.28* | -.23 | .05 | -- |

* $p < .05$

** $p < .01$

**Table 7.** ***Correlations between Trail Making (TB), Keep Track (KT), Digit Backwards (DB), and Stroop (S) tasks for Grade Three students***

| | TB | KT | DB | S |
|---|---|---|---|---|
| TB | -- | | | |
| KT | .05 | -- | | |
| DB | .26 | .20 | -- | |
| S | -.01 | -.00 | .08 | -- |

* $p < .05$

** $p < .01$

**Table 8.** ***Correlations between Trail Making (TB), Keep Track (KT), Digit Backwards (DB), and Stroop (S) tasks for Grade Five students***

| | TB | KT | DB | S |
|---|---|---|---|---|
| TB | -- | | | |
| KT | .20 | -- | | |
| DB | .08 | -.24 | -- | |
| S | -.02 | -.06 | -.01 | -- |

* $p < .05$

** $p < .01$

**Table 9.** ***Correlations between Trail Making (TB), Keep Track (KT), Digit Backwards (DB), and Stroop (S) tasks for university students***

| | TB | KT | DB | S |
|---|---|---|---|---|
| TB | -- | | | |
| KT | .10 | -- | | |
| DB | .10 | -.15 | -- | |
| S | -.03 | -.15 | .04 | -- |

* $p < .05$

** $p < .01$

**Table 10.** ***Multiple Regression Model***

| Predictor | B | Std. Error | *t* | *β* | Squared semi-partial correlation coefficients (%) | Squared partial correlation coefficients (%) |
|---|---|---|---|---|---|---|
| TB | -1.64 | 1.49 | -1.10 | -0.17 | 1.0 | 1.1 |
| KT | -0.72 | 1.24 | -0.59 | -0.06 | 0.3 | 0.3 |
| DB | -1.49 | 1.27 | -1.18 | -0.12 | 1.2 | 1.2 |
| SI | 0.02 | 0.02 | 1.29 | 0.12 | 1.4 | 1.4 |

*Note:* Model $R^2_{adj} = .01$, $F(5, 113) = 1.17$

**Table 11.** ***Correlations between Predictors and Difference Scores (DIFF) for FR Participants***

| | TB | KT | DB | SI | DIFF |
|---|---|---|---|---|---|
| TB | -- | | | | |
| KT | .32** | -- | | | |
| DB | .35** | .13 | -- | | |
| SI | -.27** | -.13 | -.04 | -- | |
| DIFF | -.08 | -.12 | -.07 | .15 | -- |

* $p < .01$

**Table 12.** ***Mean Proportions (SDs) Consistency for List 1 Details***

| | | CS | OM1 | OM2 | DS | IS |
|---|---|---|---|---|---|---|
| Grade 1 | F-R | 0.46 (0.33) | 0.23 (0.31) | 0.11 (0.22) | 0.02 (0.09) | 0.09 (0.16) |
| | R-R | 0.66 (0.28) | 0.15 (0.28) | 0.06 (0.16) | 0.07 (0.15) | 0.06 (0.14) |
| Grade 3 | F-R | 0.75 (0.25) | 0.13 (0.19) | 0.05 (0.14) | 0.04 (0.10) | 0.04 (0.10) |
| | R-R | 0.78 (0.29) | 0.04 (0.14) | 0.05 (0.17) | 0.02 (0.07) | 0.10 (0.20) |
| Grade 5 | F-R | 0.74 (0.24) | 0.10 (0.17) | 0.09 (0.14) | 0.02 (0.07) | 0.05 (0.10) |
| | R-R | 0.79 (0.23) | 0.09 (0.14) | 0.06 (0.10) | 0.02 (0.07) | 0.04 (0.09) |
| Adults | F-R | 0.83 (0.21) | 0.09 (0.16) | 0.04 (0.10) | 0.02 (0.05) | 0.01 (0.05) |
| | R-R | 0.85 (0.14) | 0.07 (0.12) | 0.04 (0.11) | 0.01 (0.03) | 0.03 (0.07) |

*Note:* Abbreviations are as follows: Consistent (CS), Omission First Recall (OM1), Omission Second Recall (OM2), Decreased Specificity (DS), Increased Specificity (IS)

**Table 13.** ***Mean Proportions (SDs) Consistency for List 2 Details***

| | | CS | OM1 | OM2 | DS | IS |
|---|---|---|---|---|---|---|
| Grade 1 | F-R | 0.72 (0.37) | 0.08 (0.19) | 0.04 (0.12) | 0.03 (0.13) | 0.04 (0.12) |
| | R-R | 0.49 (0.40) | 0.14 (0.28) | 0.11 (0.22) | 0.08 (0.18) | 0.11 (0.27) |
| Grade 3 | F-R | 0.69 (0.34) | 0.18 (0.30) | 0.03 (0.07) | 0.00 (0.00) | 0.11 (0.20) |
| | R-R | 0.63 (0.23) | 0.17 (0.20) | 0.06 (0.12) | 0.08 (0.19) | 0.06 (0.13) |
| Grade 5 | F-R | 0.70 (0.21) | 0.19 (0.22) | 0.07 (0.12) | 0.00 (0.00) | 0.05 (0.11) |
| | R-R | 0.83 (0.22) | 0.08 (0.13) | 0.00 (0.00) | 0.04 (0.11) | 0.04 (0.10) |
| Adults | F-R | 0.70 (0.28) | .09 (0.13) | 0.15 (0.20) | 0.03 (0.13) | 0.03 (.09) |
| | R-R | 0.78 (0.26) | 0.08 (0.11) | 0.04 (0.13) | 0.01 (0.04) | 0.03 (0.07) |

*Note:* Abbreviations are as follows: Consistent (CS), Omission First Recall (OM1), Omission Second Recall (OM2), Decreased Specificity (DS), Increased Specificity (IS)

# APPENDICES

## *Appendix A. Example of first interview*

**Version 2(1)**

Participant #:________________ Group Size: ____________ Gender: M / F
Interviewer:__________________ Interview Date: _________ Leader: Su / Lin

**Keep Track Task**

Practice 1 Clothing: ____________________ Insect: ______________________

Practice 2 Clothing: ____________________ Furniture: ___________________

*Test Trials*

1: An: __________ Sh: ___________ W: ___________

2: Fr: ___________ Sh: ___________ Sp: __________

3: An: __________ Sh: ___________ W: ___________ Sp: ___________

4: An: __________ Sh: ___________ W: ___________ Fr: ____________

**Making Trails**

Trail A: Time ______ (min): ______ (sec) Number of Errors: ________
Trail B: Time ______ (min): ______ (sec) Number of Errors: ________

**Free Recall**

*I wasn't here when Su/Lin came to visit, so I don't know what happened. Can you tell me everything you can remember about Bowtie Playtime. Ask* ***three non-specific prompts*** *like, "What else can you tell me about the time Su/Lin came to visit?" and "What else can you remember?" and "Can you think of anything else that happens?"*

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

## *Appendix B. Example of second interview*

**Version 2(1)**

Participant #:________________ Group Size: ____________ Gender: M / F

Interviewer:__________________ Interview Date: __________ Leader: Su / Lin

*Be sure that the participant # is clearly recorded.*

**Free Recall**

*I wasn't here when Su/Lin came to visit, so I don't know what happened. Can you tell me everything you can remember about Bowtie Playtime. Ask* ***three non-specific prompts*** *like, "What else can you tell me about the time Su/Lin came to visit?" and "What else can you remember?" and "Can you think of anything else that happens?"*

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

**Recognition**

*"You are doing great! For these questions, sometimes the right answer is 'yes' and sometimes the right answer is 'no'. If you can't remember the answer it is ok to say 'I do not know'. Sometimes I might ask you about something you just told me about, but that does not mean your last answer was wrong. It just means my teacher asked me to ask you all these questions, even if it seems like you already told me the answer, OK?"*

| | |
|---|---|
| 1. Did the experimenter ask you to think about pizza? | Y / N / DK |
| 2. Was your lucky charm a 4-leaf clover? | Y / N / DK |
| 3. Was your special helper Big Bird? | Y / N / DK |
| 4. Did you drink apple juice? | Y / N / DK |
| 5. Were you told flowers bloom best at four o'clock? | Y / N / DK |
| 6. Did you hold a starfish? | Y / N / DK |
| 7. Did you perform one of the activities at a graduation party? | Y / N / DK |
| 8. Did you wear a police officer's hat? | Y / N / DK |
| 9. Did you wear an orange apron? | Y / N / DK |
| 10. Were you paid $500 for your artwork? | Y / N / DK |
| 11. Were you told flowers bloom best at 10 o'clock? | Y / N / DK |
| 12. Did you wear a purple apron? | Y / N / DK |
| 13. Did you make trees out of clay? | Y / N / DK |
| 14. Did you learn the disappearing ball magic trick? | Y / N / DK |
| 15. Did you listen to drums playing? | Y / N / DK |
| 16. Did you listen to a piano playing? | Y / N / DK |
| 17. Did you perform one of the activities at a birthday party? | Y / N / DK |
| 18. Did you drink milk? | Y / N / DK |

19. Did you do sit-ups Y / N / DK

20. Did you sing Row, Row, Row Your Boat? Y / N / DK

21. Did you learn the falling vase magic trick? Y / N / DK

22. Did you fold a paper kite? Y / N / DK

23. Did you pretend to be Taylor? Y / N / DK

24. Did you pretend to be Alex? Y / N / DK

25. Did you colour a sticker of an airplane? Y / N / DK

26. Did you fold a paper book? Y / N / DK

27. Did you do jumping jacks? Y / N / DK

28. Did you say the code words Silly Putty? Y / N / DK

29. Were you paid $20 for your artwork? Y / N / DK

30. Did you wear a ring? Y / N / DK

31. Did you say the code words Silly String? Y / N / DK

32. Did you wear a straw hat? Y / N / DK

33. Did you colour a sticker of a car? Y / N / DK

34. Was your special helper Elmo? Y / N / DK

35. Was your lucky charm a horseshoe? Y / N / DK

36. Did the experimenter ask you to think about spaghetti? Y / N / DK

37. Did you make flowers out of clay? Y / N / DK

38. Did you hold a seashell? Y / N / DK

39. Did you sing Mary Had A Little Lamb? Y / N / DK

40. Did you wear a glove? Y / N / DK

**Digits Backwards**

| | | |
|---|---|---|
| 1. 1–4 | 6. 3–7–4–10 | 11. 6–3–9–4–10–1–8 |
| 2. 6–2 | 7. 4–8–9–1–3 | 12. 2–5–3–6–10–1–4–9 |
| 3. 1–8–5 | 8. 1–9–4–8–5 | 13. 1–6–5–9–8–3–6–4–8 |
| 4. 8–1–4 | 9. 9–6–4–8–10–1 | |
| 5. 9–5–1–8 | 10. 5–2–9–4–8–3 | |

**Stroop**

*Baseline* Column: ________ Row: ________ Cycle: 1st or 2nd

*Congruent* Column: ________ Row: ________ Cycle: 1st or 2nd

*Incongruent* Column: ________ Row: ________ Cycle: 1st or 2nd

| | | | | |
|---|---|---|---|---|
| Blue | Red | Blue | Green | Red |
| Red | Blue | Green | Red | Blue |
| Green | Green | Red | Blue | Green |
| Blue | Red | Blue | Green | Red |
| Green | Green | Red | Red | Blue |
| Red | Blue | Green | Blue | Green |
| Green | Green | Red | Green | Red |
| Red | Red | Blue | Red | Blue |
| Blue | Blue | Green | Blue | Green |
| Red | Red | Red | Green | Blue |
| Blue | Blue | Green | Blue | Green |
| Green | Green | Blue | Red | Red |
| Red | Blue | Red | Blue | Blue |
| Green | Green | Green | Red | Green |
| Blue | Red | Blue | Green | Red |
| Green | Green | Green | Blue | Blue |
| Blue | Red | Red | Green | Red |
| Red | Blue | Blue | Red | Green |
| Green | Red | Green | Blue | Blue |
| Blue | Green | Blue | Red | Red |

# REFERENCES


Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature, 410,* 366-369.

Arffa, S. (2007). The relationship of intelligence to executive function and non-executive function measures in a sample of average, above average, and gifted youth. *Archives of Clinical Neuropsychology, 22,* 969-978.

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49,* 5-28.

Baker-Ward, L., Hess, T. M., & Flannagan, D. A. (1990). The effects of involvement on children's memory for events. *Cognitive Development, 5,* 55-69.

Bala, N. (1999). Child witnesses in the Canadian courts: Recognizing their capacities and needs. *Psychology, Public Policy, and Law, 5*, 323-354

Bala, N., Lee, J., & McNamara, E. (2001). Children as witnesses: Understanding their capacities, needs, and experiences. *Journal of Social Distress and the Homeless, 10,* 41-68.

Barner, A. J., Conway, M. A., Mayoh, L., Speyer, J., Avizmil, O., & Harris, C. B. (2007). Directed forgetting of recently recalled autobiographical memories. *Journal of Experimental Psychology: General, 136,* 301-322.

Basden, B. H., & Basden, D. R. (1998). Directed forgetting: A contrast of methods and interpretations. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 139-172). Mahwah, NJ: Lawrence Erlbaum

Associates.

Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 603-616.

Bill C-2, *An act to amend the Criminal Code (protection of children and other vulnerable persons) and the Canada Evidence Act,* 1st Sess., 38th Par. (Assented July 20, 2005).

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger III, & F. I. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309-330). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Bjork, E. L., & Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness and Cognition, 5,* 176-196.

Bjork, E. L., & Bjork, R. A. (2003). Intentional forgetting can increase, not decrease, residual influences of to-be-forgotten information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 524-531.

Bjork, R. A., & Geiselman, R. E. (1978). Constituent processes in the differentiation of items in memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 347-361.

Bjork, R. A., Laberge, D., & Legrand, R. (1968). The modification of short-term memory through instructions to forget. *Psychonomic Science, 10,* 55-56.

Bjorklund, D. F., Bjorklund, B. R., Brown, R. D., & Cassel, W. S. (1998). Children's susceptibility to repeated questions: How misinformation changes children's answers and their minds. *Applied Developmental Science, 2,* 99-111.

Block, R. A. (1971). Effects of instructions to forget in short-term memory. *Journal of Experimental Psychology, 89,* 1-9.

Brady, M. S., Poole, D. A., Warren, A. R., & Jones, H. R. (1999). Young children's responses to yes-no questions: Patterns and problems. *Applied Developmental Science Special Issue: New Research on Child Witnesses, 3,* 47-57.

Brainerd, C. J., & Reyna, V. F. (2002). Recollection rejection: How children edit their false memories. *Developmental Psychology, 38*, 156-172.

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

Bray, N. W., Justice, E. M., & Zahm, D. N. (1983). Two developmental transitions in selective remembering strategies. *Journal of Experimental Child Psychology, 36,* 43-55.

Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior, 26,* 353-364.

Brewer, N., & Hupfeld, R. M. (2004). Effects of testimonial inconsistencies and witness group identity on mock-juror judgments. *Journal of Applied Social Psychology, 34,* 493-513.

Brewer, N., Potter, R., Fisher, R. P., Bond, N., & Luszcz, M. A. (1999). Beliefs and data on the relationship between consistency and accuracy of eyewitness testimony. *Applied Cognitive Psychology, 13,* 297-313.

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

593.

Bruck, M., Ceci, S. J., & Hembrooke, H. (2001). Reliability and credibility of young children's reports: From research to policy and practice. In R. Bull (Ed.), *Children and the law: The essential readings* (pp. 87-123). Malden, M.A., US: Blackwell Publishing.

Carlson, S. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28,* 595-616.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72,* 1032-1053.

Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113,* 403-439.

Ceci, S. J., & Bruck, M. (1995). The effects of repeated questioning. In S. J. Ceci & M. Bruck (Eds.), *Jeopardy in the courtroom: A scientific analysis of children's testimony* (pp. 107-125). Washington, D. C., US: American Psychological Association.

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

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

Conway, M. A., Harries, K., Noyes, J., Racsma'ny, M., & Frankish, C. R. (2000). The disruption and dissolution of directed forgetting: Inhibitory control of memory.

*Journal of Memory and Language, 43,* 409-430.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28,* 687-706.

Denckla, M. B. (1994). Measurement of executive function. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 117-142). Baltimore: P. H. Brookes Publishing.

Dent, H. R., & Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology, 18,* 41-51.

Elmes, D. G., Adams, C., & Roediger, H. L. (1970). Cued forgetting in short-term memory: Response selection. *Journal of Experimental Psychology, 86,* 103-107.

Epstein, M. A., & Bottoms, B. L. (2002). Explaining the forgetting and recovery of abuse and trauma memories: Possible mechanisms. *Child Maltreatment, 7,* 210-225.

Espy, K. A., & Cwik, M. F. (2004). The development of a Trail Making test in young children: The TRAILS-P. *The Clinical Neuropsychologist, 18,* 411-422.

Fisher, R. P., & Cutler, B. L. (1995). The relation between consistency and accuracy of eyewitness testimony. In G. Davies, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology, law, and criminal justice: International developments in research and practice* (pp. 21-28). Oxford, England: Walter De Gruyter.

Fivush, R. (1997). Event memory in early childhood. In N. Cowan (Ed.), *The development of memory in childhood* (pp.139–161). U.K: Psychology Press.

Fivush, R., & Schwarzmueller, A. (1995). Say it once again: Effects of repeated questions on children's event recall. *Journal of Traumatic Stress, 8,* 555-580.

Geiselman, R. E., Bjork, R. A., & Fishman, D. L. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General, 112,* 58-72.

Geiselman, R. E., Rabow, V. E., Wachtel, S. L., & MacKinnon, D. P. (1985). Strategy control in intentional forgetting. *Human Learning, 4,* 169-178.

Geraerts, E., & McNally, R. J. (2008). Forgetting unwanted memories: Directed forgetting and thought suppression methods. *Acta Psychologica, 127,* 614-622.

Goernert, P. N. (2005). Source-monitoring accuracy across repeated tests following directed forgetting. *British Journal of Psychology, 96,* 231-247.

Golden, C. J., & Freshwater, S. M. (2002). *The Stroop color and word test: A manual for clinical and experimental uses.* Wood Dale: Stoelting Co.

Golden, C. J., Freshwater, S. M., & Golden, Z. (2003). *Stroop color and word test children's version for ages 5-14: A manual for clinical and experimental uses.* Wood Dale: Stoelting, Co.

Goodman, G. S., Bottoms, B. L., & Rudy, L. (2001). Effects of past abuse experiences on children's eyewitness memory. *Law and Human Behavior, 25,* 269-298.

Goodman, G. S., Ghetti, S., Quas, J. A., Edelstein, R. S., Alexander, K. W., Redlich, A. D., Cordon, I. M., & Jones, D. P. H. (2003). Sexual abuse: New findings relevant to the repressed-memory controversy. *Psychological Science, 14,* 113-118.

Goodman, G. S., Goldings, J. M., & Haith, M. M. (1984). Jurors' reactions to child witnesses. *Journal of Social Sciences, 40,* 139-156.

Goodman, G. S., Tobey, A. E., Batterman-Faunce, J. M., Orcutt, H., Thomas, S., Shapiro, C., & Sachsenmaier, T. (1998). Face-to-face confrontation: Effects of closed-circuit technology on children's eyewitness testimony and jurors' decisions. *Law and Human Behavior, 22,* 165-203.

Harnishfeger, K. K., & Pope, R. S. (1996). Intending to forget: The development of cognitive inhibition in directed forgetting. *Journal of Experimental Child Psychology, 62,* 292-315.

Hershkowitz, I., & Terner, A. (2007). The effects of repeated interviewing on children's forensic statements of sexual abuse. *Applied Cognitive Psychology, 21,* 1131-1143.

Howell, D. C. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Wadsworth Group.

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

Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development, 77,* 1822-1841.

Joslyn, S. L., & Oakes, M. A. (2005). Directed forgetting of autobiographical events. *Memory & Cognition, 33,* 577-587.

Koutstaal, W., & Schacter, D. L. (1997). *Intentional forgetting and voluntary thought suppression: Two potential methods for coping with childhood trauma.* Washington, DC: American Psychiatric Association.

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

Lamb, M. E. (1999). Child witnesses: Recent research on children's accounts of forensically relevant experiences. *Applied Developmental Science, 3,* 2-5.

Lehman, E. B., & Bovasso, M. (1993). Development of intentional forgetting in children. In M. L. Howe & R. Pasnak (Eds.), *Emerging themes in cognitive development volume 1: Foundations* (pp. 214-223). New York, NY, US: Springer-Verlag.

Lehman, E. B., McKinley-Pace, M., Leonard, A. M., Thompson, D., & Johns, K. (2001). Item-cued directed forgetting of related words and pictures in children and adults: Selective rehearsal versus cognitive inhibition. *Journal of General Psychology, 128,* 81-97.

Lehman, E. B., McKinley-Pace, M., Wilson, J. A., Slavsky, M. D., & Woodson, M. E. (1997).Direct and indirect measures of intentional forgetting in children and adults: Evidence for retrieval inhibition and reinstatement. *Journal of Experimental Child Psychology, 64,* 295-316.

Lehman, E. B., Srokowski, S. A., Hall, L. C., Renkey, M. E., & Cruz, C. A. (2003). Directed forgetting of related words: Evidence for the inefficient inhibition hypothesis. *Journal of General Psychology, 130,* 380-398.

Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology, 31,* 568-578.

Lindsay, D. S., & Read, J. D. (1995). "Memory work" and recovered memories of childhood sexual abuse: Scientific evidence and public, professional, and personal issues. *Psychology, Public Policy, and Law, 1,* 846-908.

London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of child sexual abuse: What does the research tell us about the ways that children tell? *Psychology, Public Policy, and Law, 11,* 194-226.

MacLeod, C. M. (1975). Long-term recognition and recall following directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory, 104,* 271-279.

MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional Forgetting: Interdisciplinary approaches* (pp. 1-57). Mahwah, NJ: Erlbaum.

MacLeod, C. M. (1999). The item and list methods of directed forgetting: Test differences and the role of demand characteristics. *Psychonomic Bulletin & Review, 6,* 123-129.

MacLeod, C. M., Dodd, M. D., Sheard, E. D., Wilson, D. E., & Bibi, U. (2003). In opposition to inhibition. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory, Vol. 43* (pp. 163-214). New York, NY, US: Elsevier Science.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Ford, R. L. (1997). On regulation and recollection: The intentional forgetting of stereotypical memories. *Journal of Personality and Social Psychology, 72,* 709-719.

Mäntylä, T., Carelli, M. G., & Forman, H. (2007). Time monitoring and executive functioning in children and adults. *Journal of Experimental Child Psychology, 96,* 1-19.

Mazzocco, M. M. M., & Kover, S. T. (2006). A longitudinal assessment of executive function skills and their association with math performance. *Child*

*Neuropsychology, 13,* 18-45.

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

Memon, A., & Vartoukian, R. (1996). The effects of repeated questioning on young children's eyewitness testimony. *British Journal of Psychology, 87,* 403-415.

Milne, R., & Bull, R. (1999). *Investigative interviewing: Psychology and practice.* Chichester, New York: Wiley.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49-100.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81,* 111-121.

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

Nelson, K., & Gruendel, J. (1986). Children's scripts. In K. Nelson (Ed.), *Event knowledge: Structure and function in development* (pp. 21-46). Hillside N.J: Erlbaum.

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

Poole, D. A., & White, L. T. (1993). Two years later: Effects of question repetition and retention interval on the eyewitness testimony of children and adults.

*Developmental Psychology, 29,* 844-853.

Poole, D. A., & White, L. T. (1995). Tell me again and again: Stability and change in the repeated testimonies of children and adults. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 24-43). Thousand Oaks, C. A., US: Sage Publications.

Posnansky, C. J. (1976). Directed forgetting among third and seventh graders. *Contemporary Educational Psychology, 1,* 247-256.

Potter, R., & Brewer, N. (1999). Perceptions of witness behaviour-accuracy relationships held by police, lawyers, and mock-jurors. *Psychiatry, Psychology, and Law, 6,* 97-103.

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.

Price, H. L., & Connolly, D. A. (2004). Event frequency and children's suggestibility: A study of cued recall responses. *Applied Cognitive Psychology, 18,* 809-821.

Price, H. L., & Connolly, D. A. (2006). BatMon II: Children's category norms for 33 categories. *Behavior Research Methods, 38,* 529-531.

Price, H. L., & Connolly, D. A. (2007). Anxious and nonanxious children's recall of a repeated or unique event. *Journal of Experimental Child Psychology, 98,* 94-112.

Price. H. L., & Connolly, D. A. (2008). Children's recall of emotionally arousing, repeated events: A review and call for further investigation. *International Journal of Law and Psychiatry, 31,* 337-346.

Quas, J., & Schaaf, J. M. (2002). Children's memories of experienced and nonexperienced events following repeated interviews. *Journal of Experimental Child Psychology, 83,* 304-338.

Rabbitt, P. (1997). Introduction: Methodologies and models in the study of executive function. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 1-38). Hove, UK: Psychology Press.

Reitan, R. M. (1992). *Trail Making Test: Manual for administration and scoring.* South Tuscon, AZ: Reitan Neuropsychology Laboratory.

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

Reynolds, C. R., & Voress, J. K. (2007). *Test of Memory and Learning: Examiner's Manual* (2nd ed.), Austin: PRO-ED, Inc.

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

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

Sahakyan, L. (2004). Destructive effects of 'forget' instructions. *Psychonomic Bulletin and Review, 11,* 555-559.

Sahakyan, L., & Delaney, P. F. (2003). Can encoding differences explain the benefits of directed forgetting in the list method paradigm? *Journal of Memory and Language, 48,* 195-206.

Sahakyan, L., & Delaney, P. F. (2005). Directed forgetting in incidental learning and recognition testing: Support for a two-factor account. *Journal of Experimental*

*Psychology: Learning, Memory, and Cognition, 31,* 789-801.

Sahakyan, L., Delaney, P. F., & Kelley, C. M. (2004). Self-evaluation as a moderating factor of strategy change in directed forgetting benefits. *Psychonomic Bulletin and Review, 11,* 131-136.

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1064-1072.

Salmon, K., & Pipe, M-E. (1997). Props and children's event reports: The impact of a 1-year delay. *Journal of Experimental Child Psychology, 65,* 261-292.

Sas, L. D. (1993). *Three years after the verdict: A longitudinal study of the social and psychological adjustment of child witnesses referred to the Child Witness Project.* London, Ont.: London Family Court Clinic Inc., Child Witness Project.

Sas, L. D., & Cunningham, A.H. (1995). *Tipping the balance to tell the secret: Public discovery of child sexual abuse.* London: London Family Court Clinic

Semmler, C., & Brewer, N. (2002). Effects of mood and emotion on juror processing and judgments. *Behavioral Sciences and the Law, 20,* 423-436.

Sikora, D. M., Haley, P., Edwards, J., & Butler, R. W. (2002). Tower of London test performance in children with poor arithmetic skills. *Developmental Neuropsychology, 21,* 243-254.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology, 59,* 745-759.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of*

*Experimental Psychology, 18,* 643-662.

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence, 35,* 427-449.

Warren, A. R., & Lane, P. (1995). Effects of timing and type of questioning on eyewitness accuracy and suggestibility. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 44-60). Thousand Oaks, C. A., US: Sage Publications.

Whitcomb, D. (1992). Legal reforms on behalf of child witnesses: Recent developments in the American courts. In H. Dent & R. Flin (Eds.) *Children as witnesses* (pp. 151-165). Oxford, England: John Wiley & Sons.

Williams, L. M. (1994). Recall of childhood trauma: A prospective study of women's memories of child sexual abuse. *Journal of Consulting and Clinical Psychology, 62,* 1167-1176.

Williams, L. M. (1995). Recovered memories of abuse in women with documented child sexual victimization histories. *Journal of Trauma Stress, 8,* 649-673.

Wilson, S. P., & Kipp, K. (1998). The development of efficient inhibition: Evidence from directed-forgetting tasks. *Developmental Review, 18,* 86-123.

Wilson, S. P., Kipp, K., & Daniels, J. (2003). Task demands and age-related differences in retrieval and response inhibition. *British Journal of Developmental Psychology, 21,* 599-613.