Investigating General Time-Based Prospective Memory in School-Aged Children using a Novel Naturalistic Paradigm

by

Karley-Dale Talbot
B.Sc. Psychology, University of Victoria, 2011

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Psychology

© Karley-Dale Talbot, 2015
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.
Supervisory Committee

Investigating General Time-Based Prospective Memory in School-Aged Children using a Novel Naturalistic Paradigm

by

Karley-Dale Talbot
B.Sc. Psychology, University of Victoria, 2011

Supervisory Committee

Dr. Ulrich Mueller, (Department of Psychology)
Co-Supervisor

Dr. Kimberly Kerns, (Department of Psychology)
Co-Supervisor

Dr. Michael Masson, (Department of Psychology)
Departmental Member
Abstract

Supervisory Committee

Dr. Ulrich Mueller, (Department of Psychology)
Co-Supervisor

Dr. Kimberly Kerns, (Department of Psychology)
Co-Supervisor

Dr. Michael Masson, (Department of Psychology)
Departmental Member

Prospective memory (PM) refers to a person’s ability to remember to do something in the future. It is a complex behaviour that is essential for the daily functioning of young and old alike. Despite its importance in everyday life, few studies have sought to examine PM in a naturalistic way and even fewer have done so using school-aged children. The current study aimed to understand a particular form of time-based PM (TBPM), general TBPM, in children through the use of a novel naturalistic paradigm. In addition, the study aimed to add to the current PM literature by including an analysis of the circumstances surrounding a child’s prospective remembering. Results demonstrated that general TBPM was not significantly related to the parent-report Prospective Retrospective Memory Questionnaire for Children (PRMQC) or to the Memory for Intentions Screening Test for Youth (MISTY). Interestingly, general TBPM was not found to significantly relate to WM either. Descriptive analyses of the qualitative data demonstrated that no trigger rehearsals were most often responsible for children’s successful PM remembering. In contrast, when children forgot to complete their PM tasks, they most often reported being too busy with other things as the reason. The current findings provide preliminary support for the existence of a new sub-type of TBPM. They also call into question the utility of using measures like the MISTY and
PRMQC to evaluate the ecological validity of new PM task paradigms. Consequently, future research should focus on validating current PM measures before using them to evaluate the ecological validity of new ones. Finally, it is also believed that the inclusion of qualitative measures assessing the contexts of PM retrieval have important implications for the effective development of future interventions for children who experience PM difficulties.

Prospective Memory, general time-based prospective memory, PRMQC, MISTY, school-aged children, ecological validity, novel naturalistic task paradigm.
# Table of Contents

Supervisory Committee ........................................................................................................ ii
Abstract .................................................................................................................................. iii
Table of Contents .................................................................................................................. v
List of Tables ........................................................................................................................... vii
List of Figures .......................................................................................................................... xiii
Acknowledgments ..................................................................................................................... xx
Dedication ................................................................................................................................. xxi

Introduction ............................................................................................................................... 1
  Retrospective Memory and Prospective Memory ................................................................. 1
  PM ......................................................................................................................................... 2
  TBPM ................................................................................................................................... 3
  General TBPM ...................................................................................................................... 4
  TBPM in Children .................................................................................................................. 7
  Measuring TBPM Using Experimental Paradigms ............................................................... 8
  Assessing the Ecological Validity of PM Measures ............................................................ 14
  Measuring TBPM Deficits .................................................................................................... 14
  Current Study ........................................................................................................................ 17
Method ..................................................................................................................................... 18
  Participants ............................................................................................................................. 18
  Materials ................................................................................................................................. 18
    The Wechsler Abbreviated Scale of Intelligence- 2nd Edition (WASI-II) ......................... 18
    Memory for Intentions Screening Test for Youth (MISTY) .............................................. 19
    Parent Report Prospective Retrospective Memory Questionnaire for Children (PRMQC) .................................................................................................................. 20
    The Working Memory Test Battery for Children (WMTB-C) ........................................... 21
  General TBPM Task ............................................................................................................ 22
  Completed Task Questionnaire ............................................................................................. 23
  Forgotten Task Questionnaire ............................................................................................... 24
  Procedure ................................................................................................................................. 25
Results ..................................................................................................................................... 28
  Developmental Trends ........................................................................................................... 28
  Evaluating the MISTY .......................................................................................................... 28
  Correlations with the General TBPM Task .......................................................................... 29
  Investigating the Role of Cue-Specificity on General TBPM Performance .................... 30
Discussion ............................................................................................................................... 32
  Results Summary .................................................................................................................. 32
  Orthogonality of General TBPM ......................................................................................... 33
  Ecological Validity of the General TBPM Task ................................................................... 35
  Evaluation of the PRMQC ................................................................................................. 38
  Evaluation of the MISTY ...................................................................................................... 40
  Relationship between General TBPM and WM ............................................................... 41
  Reasons for Remembering and Forgetting General TBPM Tasks ................................... 42
  Limitations and Future Research ...................................................................................... 43
List of Tables

Table 1. Means and Standard Deviations for Demographic, MISTY, PRMQC, General TBPM, and Working Memory Subtest Measures ................................................................. viii
Table 2. Means and Standard Deviations for Demographic, MISTY, PRMQC, General TBPM, and Working Memory Subtest Measures by Age Group ........................................ ix
Table 3. Correlation and Significance Values Associated with the Sample ......................... x
Table 4. Partial Correlation and Significance Values Associated with the Sample after Controlling for Age ........................................................................................................ xi
Table 5. Means and Standard Deviations Associated with MISTY-PM Errors .................. xii
Table 1. Means and Standard Deviations for Demographic, MISTY, PRMQC, General TBPM, and Working Memory Subtest Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>29</td>
</tr>
<tr>
<td>IQ</td>
<td>112.52 (13.97)</td>
</tr>
<tr>
<td>Percent Male</td>
<td>51.7</td>
</tr>
<tr>
<td>Age</td>
<td>9.31 (1.91)</td>
</tr>
<tr>
<td>MISTY Score</td>
<td>33.48 (11.96)</td>
</tr>
<tr>
<td>MISTY Timed Score</td>
<td>5.34 (2.21)</td>
</tr>
<tr>
<td>PRMQC Score</td>
<td>20.24 (6.11)</td>
</tr>
<tr>
<td>General TBPM Score</td>
<td>2.90 (1.65)</td>
</tr>
<tr>
<td>WMTB-C Mean Raw Score</td>
<td>18.67 (4.91)</td>
</tr>
</tbody>
</table>

Note:
1. Total score on the MISTY, out of a total of 48
2. Total score on the MISTY timed cues, out of a total of 8
3. Total score on the PM questions of the PRMQC, out of possible total of 40
4. Number of successfully completed trials on the general TBPM task, out of a total of 6.
5. Average raw score on the working memory subtests (listening recall and dot counting), out of possible total of 78

<table>
<thead>
<tr>
<th></th>
<th>7 years</th>
<th>8 years</th>
<th>9 years</th>
<th>10 years</th>
<th>11 years</th>
<th>12 years</th>
<th>13 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>6</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>FSIQ</td>
<td>111.00</td>
<td>114.57</td>
<td>113.00</td>
<td>112.75</td>
<td>113.40</td>
<td>112.00</td>
<td>107.00</td>
</tr>
<tr>
<td>MISTY PM Score</td>
<td>19.50</td>
<td>33.43</td>
<td>34.00</td>
<td>42.50</td>
<td>36.60</td>
<td>34.50</td>
<td>48.00</td>
</tr>
<tr>
<td></td>
<td>(12.69)</td>
<td>(8.20)</td>
<td>(15.10)</td>
<td>(3.32)</td>
<td>(7.47)</td>
<td>(10.67)</td>
<td>(.00)</td>
</tr>
<tr>
<td>MISTY Timed Score</td>
<td>3.17</td>
<td>5.43</td>
<td>6.00</td>
<td>6.50</td>
<td>5.80</td>
<td>4.50</td>
<td>8.00</td>
</tr>
<tr>
<td></td>
<td>(2.64)</td>
<td>(1.99)</td>
<td>(2.00)</td>
<td>(1.29)</td>
<td>(1.48)</td>
<td>(2.12)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>PRMQC-PM Score</td>
<td>18.83</td>
<td>19.00</td>
<td>18.67</td>
<td>16.75</td>
<td>15.80</td>
<td>19.50</td>
<td>13.00</td>
</tr>
<tr>
<td></td>
<td>(3.71)</td>
<td>(6.06)</td>
<td>(6.03)</td>
<td>(5.25)</td>
<td>(2.49)</td>
<td>(13.44)</td>
<td>(4.24)</td>
</tr>
<tr>
<td><em>General</em> TBPM Score</td>
<td>2.50</td>
<td>3.14</td>
<td>4.33</td>
<td>2.25</td>
<td>3.40</td>
<td>4.50</td>
<td>2.50</td>
</tr>
<tr>
<td></td>
<td>(.84)</td>
<td>(2.12)</td>
<td>(1.15)</td>
<td>(2.22)</td>
<td>(.55)</td>
<td>(2.12)</td>
<td>(2.12)</td>
</tr>
<tr>
<td>WMTB-Mean Raw</td>
<td>19.42</td>
<td>19.07</td>
<td>18.17</td>
<td>19.38</td>
<td>17.40</td>
<td>20.00</td>
<td>16.75</td>
</tr>
<tr>
<td></td>
<td>(4.35)</td>
<td>(5.09)</td>
<td>(2.93)</td>
<td>(4.50)</td>
<td>(7.61)</td>
<td>(0.00)</td>
<td>(8.13)</td>
</tr>
</tbody>
</table>
Table 3. Correlation and Significance Values Associated with the Sample.

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>WASI Full Scale IQ</th>
<th>MISTY-PM</th>
<th>MISTY-Timed</th>
<th>PRMQC</th>
<th>General TBPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>WASI Full Scale IQ</td>
<td>-.045</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MISTY-PM</td>
<td>.559**</td>
<td>.293</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MISTY-Timed</td>
<td>.431*</td>
<td>.388*</td>
<td>-.875**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRMQC</td>
<td>-.169</td>
<td>-.379*</td>
<td>-.215</td>
<td>-.424*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>General TBPM</td>
<td>.104</td>
<td>-.247</td>
<td>-.021</td>
<td>-.009</td>
<td>.162</td>
<td></td>
</tr>
<tr>
<td>WMTB-C Raw Mean</td>
<td>.704**</td>
<td>.057</td>
<td>.614**</td>
<td>.620**</td>
<td>-.408*</td>
<td>-.142</td>
</tr>
</tbody>
</table>

*significant at the $p < .05$
**significant at the $p < .01$
Table 4. Partial Correlation and Significance Values Associated with the Sample after Controlling for Age

<table>
<thead>
<tr>
<th></th>
<th>WASI Full Scale IQ</th>
<th>MISTY-PM</th>
<th>MISTY-Timed</th>
<th>PRMQC</th>
<th>General TBPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MISTY-PM</td>
<td>.342</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MISTY-Timed</td>
<td>.412*</td>
<td>.828**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRMQC</td>
<td>-.365</td>
<td>.066</td>
<td>-.153</td>
<td></td>
<td></td>
</tr>
<tr>
<td>General TBPM</td>
<td>-.151</td>
<td>.013</td>
<td>.023</td>
<td>.212</td>
<td></td>
</tr>
<tr>
<td>WMTB-C Raw Mean</td>
<td>.060</td>
<td>.090</td>
<td>.181</td>
<td>.018</td>
<td>-.021</td>
</tr>
</tbody>
</table>

*significant at the \( p < .05 \)
**significant at the \( p < .01 \)
Table 5. Means and Standard Deviations Associated with MISTY-PM Errors

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Errors</td>
<td>2.66 (2.01)</td>
</tr>
<tr>
<td>Prospective Failures</td>
<td>1.14 (1.75)</td>
</tr>
<tr>
<td>Task Substitutions</td>
<td>.34 (.67)</td>
</tr>
<tr>
<td>Loss of Content</td>
<td>.55 (.78)</td>
</tr>
<tr>
<td>Loss of Timing</td>
<td>.59 (.73)</td>
</tr>
<tr>
<td>Random</td>
<td>.03 (.186)</td>
</tr>
</tbody>
</table>

Note:
1. Total errors on the MISTY, out of a possible 8
2. Child did not respond to target
3. Child mixed up task demands (e.g., performed one task at the time another was supposed to be executed)
4. Child recognized that a task needed to be performed, but forgot what s/he needed to do
5. Completed a PM task at the wrong time (e.g., after 5 minutes instead of 10). Received a score of 1, not 0.
List of Figures

Figure 1. Accuracy in PM performance by PM task type............................... xiv
Figure 2. Relationship between general TBPM scores and MISTY TBPM scores. ....... xv
Figure 3. Percentage by circumstance type on successful general TBPM tasks.......... xvi
Figure 4. Percentage of MISTY errors by error type................................... xvii
Figure 5. Percentage of general TBPM failures by circumstance type .................. xviii
Figure 6. Percentage of participants who when they forgot their task, remembered the
task at some point during target day by frequency ........................................ xix
Figure 1. Accuracy in PM performance by PM task type.

Note:
1 Percent correct out of total PM trials only, out of 48
2 Percent of total severity score on PM trials only, out of 45
3 Percent correct out of total TBPM trials, out of 6
Figure 2. Relationship between general TBPM scores and MISTY TBPM scores.
Figure 3. Percentage by circumstance type on successful general TBPM tasks.
Figure 4. Percentage of MISTY errors by error type.
Figure 5. Percentage of general TBPM failures by circumstance type
Figure 6. Percentage of participants who when they forgot their task, remembered the task at some point during target day by frequency.
Acknowledgments

This research was supported by a grant from the Natural Sciences and Engineering Research Council. I would like to thank my co-supervisors, Drs. Kimberly Kerns and Ulrich Mueller for their insight, advice, and support throughout the development and execution of this study. I would also like to thank my friends, family, and colleagues for their help with the recruitment process. Finally, I would especially like to thank the parents and children who generously donated their time. If it were not for all of your help and support, this would not have been possible.
Dedication

I would like to dedicate this thesis to my mother. She truly has made me the woman I am today. She has taught me to be strong and persevere when things get tough. Without her unconditional love and support, I would not be where I am today. Although she will be the first to tell you that she will not understand much of this thesis, I know she is very proud - and that is all that matters.
Introduction

Prospective memory (PM) is the cognitive process associated with the control and coordination of future actions and activities (Ellis, 1996). It encompasses an individual’s ability to remember to do something at some point in the future and is essential for the successful completion of a range of everyday activities, from the simple and inconsequential to the extremely important. To illustrate, a person uses her PM ability to remember to buy milk on the way home or to remember to give her son his insulin on time. It is because of this daily reliance on PM that researchers have shown considerable interest in trying to understand the mechanisms underlying its successful execution. In particular, this study aims to understand a particular form of PM using a newly developed naturalistic paradigm. In the first section PM will be distinguished from the more commonly known retrospective memory (RM). Next, general and specific forms of time-based PM will be differentiated. Following this, several experimental weaknesses in current laboratory-based PM tasks will be identified.

Retrospective Memory and Prospective Memory

Though both are forms of memory, retrospective memory (RM) and PM differ in a number of ways. The most significant difference is that unlike RM, PM contains a future-oriented component that requires one to remember to act upon on one’s own. Additionally, experimental PM tasks differ from experimental RM tasks. In particular, PM tasks typically consist of: (1) a delay between the formation of an intention and its performance, (2) a lack of an explicit reminder to perform the intention, and (3) a need to disrupt an ongoing task in order to perform the intended action (Rendell, Vella, Kliegel, & Terrett, 2009). Interestingly, it has also been found that the successful execution of PM
relies more heavily on the more frontal lobe dependent executive functions than does RM. Specifically, PM has been found to rely on the executive functions of attention, planning, and inhibitory control for its successful completion (Groot, Wilson, Evans, & Watson, 2001; Kliegel, Martin, McDaniel, & Einstein, 2001; Mackinlay, Kliegel, & Mäntylä, 2009; Shum, Cahill, Hohaus, O’Gorman, & Chan, 2012).

**PM**

According to Ellis (1996), there are five general phases in the successful execution of PM: (1) formation and encoding, (2) a retention interval, (3) a performance interval, (4) initiation and execution of the intended action, and (5) evaluation of the outcome. Phase 1 is acknowledged as the RM component of a PM task as it involves long-term storage of an intention which requires the use of encoding. Consequently, phases 2 to 5 constitute the prospective component. In the first phase, the focus is on forming an intent, retaining the action, and encoding the retrieval context. Put more simply, the first phase involves deciding what one wants to do and in what context one wants to do it. The second and third phases refer to the delay between encoding and retrieval and the point when the intended action should be retrieved respectively. It is important to note that the retention and performance interval durations may vary considerably and that an associated memory for action can be retrieved at any point during these two phases. That being said, in order for one to successfully perform PM, the delayed intention must be elicited during a performance interval when the retrieval context matches that which was encoded in phase 1. At this point, phases 4 and 5 can take place, completing the action.

According to Kvavilashvili and Ellis (1996), PM can be separated into several distinct subtypes depending on the context being used to trigger the retrieval of an
intention. That being said, the majority of the PM literature has focused primarily on the
time- and event-based subtypes of PM. Event-based PM (EBPM) has been defined by
Ellis (1996) as a memory for action in response to a specific cue. This happens when one
remembers to pass a note (action) to a co-worker after seeing him (cue) in the hallway. In
contrast, time-based PM (TBPM) has been defined as a memory for action “at a certain
time or after a certain amount of time has elapsed” (Einstein and McDaniel, 1996; p.
119). This means that TBPM occurs when one remembers to go to a doctor’s
appointment at 3pm or when one remembers to call a friend again sometime in the
afternoon when he will return. This paper focuses on TBPM.

**TBPM**

For TBPM, the appropriateness of the action is determined by the passage of time
rather than by the occurrence of an event (Einstein & McDaniel, 1996). This means that a
person must rely on internal or self-generated monitoring and retrieval processes rather
than external cues in order to successfully execute a TBPM task. It is because of this that
TBPM tasks are hypothesized to be more difficult than EBPM tasks (Einstein &
McDaniel, 1996). A major line of research has investigated the psychological processes
underlying successful PM. One process that has received considerable attention is
executive function, which refers to the higher-order mental processes such as planning,
problem solving, self-monitoring, and working memory (WM). While self-initiated
monitoring would seem to form the basis for the association between TBPM and
executive function (Guerten, Lejeune, & Meulemans, 2014), several studies conducted to
investigate this association have yielded conflicting results. For example, a study by
Mackinlay, Kliegel, and Mäntylä (2009) looked at several hypothesized predictors of
TBPM in a group of typically-developing school-aged children and found that the executive functions of planning, task switching, and WM were the most important predictors of laboratory-based TBPM performance. Similarly, Mahy and Moses (2011) found that WM, but not inhibitory control, predicted TBPM once the effect of age was controlled for. Interestingly, Kerns (2000) found both WM and inhibitory control to be related to TBPM performance.

Harris and Wilkins (1982) proposed the Test-Wait-Test-Exit (TWTE) model of TBPM that evolved from previous work by Miller, Galanter, and Pribram in 1960 (as outlined in Einstein & McDaniel, 1996, and McDaniel & Einstein, 2007). The TWTE model describes the process of TBPM as a mechanism that requires continuous monitoring or checking of the current status of a task until the appropriate time to respond is reached. According to the authors, it is at this point that the target action is then initiated and executed. Therefore, this model suggests that successful remembering is critically dependent on monitoring. Monitoring refers to the act of checking the time during the “critical period”, or “appropriate time to respond”, and is believed to be mainly self-initiated. This also suggests that there may be an association between TBPM and time estimation, a process that is also thought to rely on executive functions (Aberle & Kliegel, 2010; Kerns, 2000; Zinke, Altgassen, Mackinlay, Rizzo, Drechsler, & Kliegel, 2010).

**General TBPM**

According to the TWTE model, successful remembering requires one to check the clock at the intended time (Einstein & McDaniel, 1996, p. 130). However, what happens when the appropriate moment to respond is not specified, but is a time range. Previous
research by Ellis (1996) has defined these types of intentions as *Step* intentions (p.13).

However, for the purposes of this paper, the term ‘general TBPM’ will refer to the case when retrieval time is some time range or a more generalized period of time for the intended action (i.e., sometime on Tuesday). This is in contrast to *specific* TBPM that refers to memory for action at a particular time (i.e., 9am). Although *specific* TBPM has received considerable attention by researchers, there are very few studies that examine *general* TBPM directly.

A literature review revealed few studies that have used a *general* rather than *specific* TBPM measure. Of these, three will be outlined below. Cook, Marsh, and Hicks (2005), investigated the effect of expected context on time-based and event-based PM performance. On the TBPM task, the participants were asked to press a key after 6 minutes, but before 7 minutes had elapsed on a clock. This task required self-initiated checking. Because the participants were able to press the button at any time within a 60-second interval, the response time for this task was not specific and could therefore be considered a *general* TBPM task.

Similar to Cook and colleagues’ (2005) task, Kerns (2000), and Talbot and Kerns (2013) used a TBPM task that provided participants with a 10-15 second time range during which they could perform the PM task (e.g., refill the fuel tank of their car/spaceship during video game play). Because the response times for these tasks were not dependant on the participants refuelling at a specified clock time, these tasks could also be considered *general* TBPM tasks, as per the definition provided above. That said, these trials were very close together, potentially facilitating learning and therefore, possibly decreasing the sensitivity of the task across trials (as monitoring may have
become ‘habit’ for the participants). The current study aims to develop a TBPM task using longer time frames during which the intended action may take place in an attempt to better represent individual differences in participants’ PM abilities.

Since general TBPM allows for a time range for responding instead of a specific time, it has been suggested to be easier than specific TBPM (Ellis, 1996, p. 14; see also Kvavilashvili & Ellis, 1996). This is believed to be because a general TBPM task offers numerous opportunities to perform the intended action (at any time within a specified time range). This is in contrast to a specific TBPM task that offers a single (e.g., at 3pm) opportunity. On the other hand, it is possible that the potential advantage believed to be associated with general TBPM may be offset by the nature of the target retrieval context initially encoded in memory. That is, the target context (e.g., the appropriate time to respond) may actually reflect many retrieval contexts that are in some ways restricted by a person’s daily activities (Einstein & McDaniel, 1996). For example, when a person encodes an intention to buy toothpaste that day, the target retrieval context (e.g., a time period long enough to go to the store and buy the toothpaste) may be restricted to a person’s lunch break or drive home from work, thus severely limiting the person’s number of opportunities to complete the task. To date there has been no empirical research investigating the effect of a more general time range on TBPM performance and it is unknown whether specific or general TBPM tasks are easier to perform.

Another factor proposed to affect the performance of general TBPM is the habituation of the intended action (i.e., the extent to which the intended action is routinely performed). Researchers have located tasks on a continuum with the two poles: habituated and episodic (Ellis, 1996; Kvavilashvili & Ellis, 1996). Episodic tasks are
TBPM tasks that are infrequently performed such as buying toothpaste on the way home. Habitual tasks, however, are tasks that are routinely performed, like brushing one’s teeth before going to bed. According to Ellis (1996), the extent of habituation of a task will have differential effects on the successful performance of the PM task. This means that habitual tasks are more likely to be remembered over episodic tasks because they are more strongly encoded in memory (e.g., habits). That said, the level of motivation to perform an episodic task (likely tied to reward or punishment based on performance of the task) may counteract this potential disadvantage (see below for a discussion). Despite the lack of studies investigating this general TPBM, it comes into play regularly in everyday life and is therefore an important topic for research.

**TBPM in Children**

Extant research agrees that TBPM follows a developmental trajectory with older children tending to outperform younger children on laboratory-based measures (Ceci & Bronfenbrenner, 1985; Kerns, 2000; Kerns & Price, 2001; Mackinlay et al., 2009; Voigt, Aberle, Schönfeld, Kliegel, personal communication). In addition, studies have shown that children begin to develop the ability to perform TBPM by the age of 5 years and that this ability continues to develop until about the age of 13 years when it is hypothesized to have reached an adult level of performance (Aberle & Kliegel, 2010; Kliegel & Mackinlay, & Jäger, 2008; McDaniel & Einstein, 2007). However, the majority of studies have investigated only the development of TBPM in a laboratory setting. The current study will instead aim to determine if TBPM follows a similar developmental trajectory in more naturalistic settings.
Measuring TBPM Using Experimental Paradigms

The vast majority of the research currently available on TBPM has looked at specific TBPM in a laboratory setting using variations of the general experimental paradigm put forth by McDaniel and Einstein (2006). McDaniel and Einstein (2006) described a general approach for measuring TBPM in a controlled laboratory setting and outline 4 steps to follow to ensure that specific TBPM is being accurately measured: (1) keep participants busy with an ongoing task (e.g., pleasantness rating), (2) at the beginning of the ongoing task, ask participants to perform another task (e.g., key press) at a specific time during the experiment, (3) give a filler task between the initial task instructions and the commencement of the ongoing activity to reduce the likelihood of the intention being maintained in WM, and (4) measure performance by the proportion of trials in which participants remember to execute the PM task. They argued that this general approach provides an ecologically valid measurement of specific TBPM that can be conducted in a controlled laboratory setting. They felt that by embedding the PM task within an ongoing task, the participant’s specific TBPM performance would be an accurate measure of their everyday performance.

As is the case with many research paradigms, variants of this experimental design have been introduced. With respect to specific TBPM, the ongoing task and the PM task can vary greatly across studies while still following the general outline proposed by McDaniel and Einstein (2006). To illustrate, some studies have used measures of cognitive function as their ongoing task (Kliegel et al., 2005; Mäntylä, Del Missier, & Nilsson, 2009; Mackinlay, Kliegel, & Mäntylä, 2009) while others have used word rating tasks (Labelle, Graf, & Grondin, 2009; Nater et al., 2006). Additionally, some researchers have used simple, child-friendly concentration games (Aberle & Kliegel, 2010) or have...
had participants listen to stories (Mäntylä, Del Missier, & Nilsson, 2009). Interestingly, the types of PM tasks generally employed by researchers do not tend to vary as much as the ongoing tasks and usually consist of pressing a computer key at a specific point in time or after a specified time interval (Khan, Sharma, & Dixit, 2008; Labelle, Graf, & Grondin, 2009; Mackinlay, Kliegel, & Mäntylä, 2009). That said, the duration and frequency of the key presses do tend to vary across studies and can range from approximately every minute for 5 minutes (Kerns, 2000) to every 15 minutes (Khan, Sharma, & Dixit, 2008). In addition, determining when it is the right time to press the key is usually facilitated by self-initiated checking of either a visible clock or gauge that can be accessed with a keystroke.

Although laboratory-based measures of specific TBPM follow the guidelines set out by McDaniel and Einstein (2006) and are conducted in a more controlled setting, results may only loosely reflect participants’ everyday PM abilities (Rose, Rendell, McDaniel, Aberle, & Kliegel, 2010). This is believed to be because the majority of current laboratory-based PM tasks are either not representative of typical daily PM tasks or the constrained laboratory environment may unnecessarily influence a participant’s PM performance (Sbordone, 1996). Consequently, it is difficult to accurately interpret and predict future PM functioning (Goldstein, 1996; Sbordone, 1996). For this reason, researchers have begun to investigate the ecological validity of current laboratory-based PM measures and develop new, more naturalistic paradigms to more accurately measure daily PM functioning.

To date, there are few studies investigating PM using naturalistic paradigms. Of the naturalistic studies available, there is only a handful that investigates TBPM specifically.
A particularly interesting study conducted by Rose and colleagues (2010) attempted to create a more ecologically valid task with young and older adults using a contextual board game, called the “Virtual Week Task”, meant to mirror the real life demands of PM. In this task participants were asked to play a board game that simulated going through the course of a day for 5 consecutive days. During this game, participants were asked to remember to “perform” a number of PM tasks at different points in “game time” or as a result of the occurrence of certain “game daily” events. The game incorporated 50 PM cues – 10 per “day” across the 5 “days”. These PM cues varied in frequency of occurrence (i.e., repeated vs. non-repeated) and included two time-check tasks per “day”. For example, the authors had the participants remember to “take” an antibiotic at each ‘game’ breakfast and dinner events as a type of repeated task. Conversely, the researchers included a task of having to phone a bank at the games ‘noon time’ to arrange an appointment as a type of non-repeating task. Although participants were told to carry out these tasks at the appropriate times, they were not asked to perform them physically. Instead they were to perform the task by choosing it from a list of PM tasks and distractors. Similar to the repeated tasks, the time-check tasks required the participant to remember to take a “lung capacity” test (breathe in and out of a spirometer that records the amount and rate of air that one breathes in over time) when the clock at the center of the board, read 2 minutes and 15 seconds, and after the start of each “day”. Unlike the other task types, participants were asked to physically perform these tasks.

Although the “Virtual Week” task attempted to produce a much more naturalistic paradigm that required task switching and WM, the study was still conducted in a controlled laboratory environment over a relatively short time duration, and required
participants to commit to memory both the repeated and time-check events on three occasions before allowing them to start the game. Additionally, although the PM tasks included a variety of everyday situations, the time-check tasks did not appear to be representative of a real-life scenario as it is unlikely that individuals would need to test their lung capacity on a regular basis or at such a specific times. This suggests that this task could have been done a minute before or after the specified time without significant consequences, thus better representing a general TBPM task in a natural setting.

Finally, the fact that the participants “lived” 5 days in one sitting where they were asked to remember 50 different tasks is also not representative of their everyday PM tasks. Firstly, it is unusual that an individual would need to prospectively remember 10 different things in one day without the use of external reminders like lists or alarms. Secondly, it is unlikely that a person would include a distractor in his list of things to do that day as it would add an additional memory component to the task making it harder to carry out. Lastly, people do not generally submit repeated tasks to memory before carrying them out to ensure that they become habitual. Instead, these tasks generally become habitual gradually over an extended period of time. As such, the ecological validity of this task is compromised.

Another quasi-naturalistic study conducted by Walsh, Martin, and Courage (2013) used a virtual reality computer-based task called the “Shopping Trip” task in an effort to evaluate age differences in PM as a function of increasing cue retrieval specificity in preschool-aged children. In the first experiment, participants were asked to remember a shopping list consisting of pictures of stores to visit and particular things to buy or do at each store (primary visually-cued EBPM task). In addition, they were also shown a
picture of Elmo and were told “touch him” on the screen as quickly as possible whenever they saw him appear (secondary visually-cued EBPM task). Finally, the participants were asked to remind the researcher to give them stickers at the end of the game (tertiary incidental activity-based PM task). The participants were then asked to help the researcher complete a puzzle as a distractor task.

Similar to the Rose et al. (2010) study, the Walsh et al. (2013) study is better considered to be quasi-naturalistic in that it was also conducted in a controlled laboratory environment. In addition, the preschoolers were asked to complete a shopping trip, remembering at what store they were to stop, and what they were to do or buy in an effort to emulate real-life “play shopping”. The authors of the study believed this to be a “common everyday activity” for preschoolers. With the exception of children ‘pretending’ to shop, children are unlikely to be responsible for remembering which stores their parents will have to shop at and/or what items their parents need to buy. Further, children at this age are unlikely to be paying attention to stores that are not of particular interest to them (e.g., jewelry store vs. toy store) when they go shopping with their parents. Similarly, touching a highly salient and frequent visual cue (e.g., Elmo) is not likely to provide an accurate representation of participants’ every day visually-cued EBPM abilities. For these reasons, this task should not be considered an accurate representation of an everyday PM activity that a preschooler would be required to carry out. Therefore, the results should not be interpreted as being representative of preschoolers’ everyday EBPM ability. It is for this reason that the current study will introduce a more age-appropriate naturalistic task conducted over a series of weeks, outside of the laboratory.
In contrast to the quasi-naturalistic studies outlined above, another type of naturalistic study involving adult participants had individuals conduct particular tasks while they were at their own homes. For example, Kvavilashvili and Fisher (2007) had participants remember to phone the researcher at a pre-arranged time within a one-week interval. Within that week, the researchers also had participants record in detail the occasions when they had thought about the intention. Similarly, Szarras and Niedźwieńska (2011) had their participants generate a list of tasks that they were planning to perform over the course of the next 10 days. During this time, the participants were asked to record the details about each time they thought about any of the tasks they had previously listed. In addition, they were asked to rate the nature of any triggers or cues that prompted their memory rehearsals as either (1) an accidental rehearsal, (2) a self-initiated rehearsal, or (3) a no-trigger rehearsal. An ‘accidental rehearsal’ was defined as any instance in which the participant’s memory was accidentally triggered either externally by visual stimuli or internally by cognitive stimuli. A ‘self-initiated rehearsal’ was chosen by participants when they had actively tried to remember what it was they were supposed to do. In contrast, a ‘no-trigger’ rehearsal would be chosen if the participants believed that the memory for action had just popped into their mind without having been triggered by anything. The benefit of experiments like these is that they allow the researcher to get a better sense of the types of delayed intentions people form and how their memories for action are triggered in everyday life. They have the capacity to provide richer information about why an action is occurring than many of the laboratory-based experiments currently in use. The current study aims to investigate the performance of
general TBPM in children using a naturalistic paradigm that provides information about the context of memory retrieval.

**Assessing the Ecological Validity of PM Measures**

According to Sbordone (1996), ecological validity refers to the relationship between psychological test scores and a person’s present and future functioning in real-world settings. It is an exceedingly important property to evaluate as the main goal of psychological tests is to understand a person’s difficulties and to determine how that might translate into everyday life (Sbordone, 1996). Despite its importance, the predictive relationship between performance on PM test measures and real-world behaviour has been largely ignored by PM researchers. Consequently, the relationship between current PM measures and everyday PM functioning is unclear. It is for this reason that a major aim of the current study was to evaluate the ecological validity of the general TBPM task.

According to Franzen and Wilhelm (1996) there are two general aspects of ecological validity: (1) verisimilitude and (2) veridicality. Consideration of verisimilitude occurs during the design and development of a neuropsychological test measure. It refers to the similarity of the data collection method to the tasks and skills that are required of persons during their everyday lives. Conversely, the veridicality of a test measure is evaluated once the task has been designed. It refers to the extent to which the results of a test measure reflect or predict the behaviour in the open environment.

**Measuring TBPM Deficits**

Although the combination of laboratory and naturalistic research has aided our understanding of PM, assessment of PM is also important within the clinical context,
where individuals with deficits in PM require identification for remediation/interventions. Currently, the most popular screening measure of PM ability in adults is the Memory for Intentions Screening Test (MIST™). The MIST was designed to be a relatively brief clinical measure of PM to be used with adult populations (e.g., dementia, traumatic brain injury) who frequently experience deficits in PM (Raskin, 2009). This measure includes both event- and time-based PM cues and has items with varying delay periods and response types. It has also been found to be a sensitive and specific measure in clinical groups, with demonstrated adequate psychometric properties.

In 2009, Mills and colleagues developed the Memory for Intentions Screening Test for Youth (MISTY) to measure PM deficits in children, based on the MIST™ protocol. The MISTY contains eight different tasks that vary in response type, delay period, and cue-type, but has shorter time delays and more child-appropriate PM tasks. In a preliminary study, it was believed that the MISTY measured PM in children effectively (Mills et al., personal communication). That said, it is important to note that the MISTY remains subject to all of the same criticisms associated with the laboratory-based PM measures previously identified.

Another approach to the naturalistic measurement of PM for an individual is to obtain reports of PM performance via report from others who may have experience with an individual’s ability to adequately complete future intentions. With this goal in mind, Talbot and Kerns (2013) modified an adult-report Prospective Retrospective Memory Questionnaire, designed by Smith and colleagues (2000), to develop the Prospective Retrospective Memory Questionnaire for Children (PRMQC). The PRMQC is a brief (16-item) parent-report questionnaire that asks parents to respond to simple statements
about the frequency of their child’s RM and PM failures. For example, one question asks: “Does your child fail to do something he/she was supposed to do a few minutes later even though it’s there in front of him/her, like turning off the TV or Gameboy or picking up their backpack before heading to school?” (PM question). The PRMQC was used in a study with children with ADHD and was found to relate to laboratory-based measures of PM and parent reports of ADHD symptom severity (Talbot & Kerns, 2013). Likewise, Kliegel and Jäger (2007) developed a different modified version of the PRMQ to be used with preschoolers. Their version also found that parents and caregivers were able to adequately rate their children’s RM and PM difficulties. Taken together, both adapted versions of the PRMQ appear to be valid and sensitive measures of everyday PM errors in children, despite their reliance on parent-report, which can be unduly influenced by parents’ worries about their child’s behaviours. That said, one important thing to note is that the PRMQC scores in both of the described studies were not broken down by EB or TBPM (specific or general) items and as such, provide only a generalized measure of PM difficulties in children.

Assessment measures like the PRMQ and PRMQC are useful as they can potentially provide an objective measure of PM ability that can then be used to ascertain the ecological validity of laboratory-based measures of PM. Alternatively, measures like the MIST and MISTY have the potential to aid in the diagnosis of particular PM deficits in individuals with deficits in these abilities, which can then be used to dictate areas for intervention and/or rehabilitation. As such, the current study will investigate the relations between scores on the MISTY, PRMQC, and the newly devised naturalistic general TBPM measure.
**Current Study**

Interest in the role of TBPM in children’s daily functioning has increased substantially in the last 15 years. Yet, no research has investigated *general* TBPM performance using a naturalistic paradigm. Likewise, whereas there is some evidence showing that the MISTY and PRMQC are sensitive and specific measures of children’s laboratory-measured *specific* PM, their association with a naturalistic measure of *general* TBPM has not been examined. The primary goal of the present study was to investigate *general* TBPM in an ecologically valid way. The current study also aimed to: (1) investigate the role of cue-specificity on *general* TBPM performance; (2) Compare the PRMQC to both the MISTY and a naturalistic measure of *general* TBPM; (3) and provide further evidence of the sensitivity and psychometric properties of both the MISTY and the PRMQC. Finally, because WM has consistently been found to relate to laboratory-based measures of TBPM, the current study also aimed to investigate whether the novel *general* TBPM task would be associated with WM measures as well.
Method

Participants

29 children (15 males, 14 females), predominately of European heritage (66%), aged between 7 to 13 years participated in this study. These children were recruited through the use of departmental listerves, word of mouth, school and community postings, and social media sites (i.e. Facebook, kidsinvictoria.com). As a thank you for participation, each participant was given the chance to win two tickets to the movie theatre that were awarded to a single participant at the completion of the study.

In order to be enrolled in the study, a parent of each participant was required to actively participate in the study with his/her child. In addition, screening criteria included participants having an IQ higher than 85, no pre-existing neurologic condition or brain injuries, or a learning disability, to ensure assessment of general TBPM in a typically developing sample of children in this age range (see Table 1). All participants’ parents signed Human Research Ethics Board-approved informed consent forms and children gave verbal assent prior to commencing the study.

Materials

The Wechsler Abbreviated Scale of Intelligence- 2nd Edition (WASI-II)

This published standardized measure was developed by Wechsler (2011) to provide a brief, reliable measure of cognitive ability in individuals aged 6 to 90. It can be used to estimate a range of full-scale IQ (FSIQ) scores. For the current study, the researcher used this measure to screen for the existence of possible intellectual disabilities. The measure included four separate subtests (e.g., Block Design, Vocabulary, Matrix Reasoning, and Similarities). During this task participants were asked to construct block designs based on
pictures, define words, choose pictures that best fit puzzles with a missing piece, and explain how two words are alike (e.g. “how are ‘love’ and ‘hate’ alike?”).

**Memory for Intentions Screening Test for Youth (MISTY)**

The MISTY, an approximately 20-minute paper-and-pencil test is an adaptation for youth based on the adult Memory for Intentions Test™ (MIST™; Raskin & Buckheit, 1998). For this measure, researcher and participant are seated at a table with a digital clock in view. During the task, participants perform eight separate PM tasks, four each of time- and event-based trials, consisting of both trials requiring a verbal response or an action. For example, a time-based verbal response trial might be, “In 2 minutes, at 3:05, tell me to do my homework” whereas, “When I hand you a blue marker, draw a house” is an example of an action based EBPM trial. Each of the trials is then assigned to three different scores of interest: cue type (event or time), time delay (2 or 10 minutes), and response type (action or verbal). After the participants have been presented with their tasks, they are asked to perform an ongoing word search game. Following completion of the final PM trial, the participants were asked to answer eight multiple-choice questions to assess RM of the PM cues for the task. Finally, type of error was assigned for each failed trial, coded as either a (1) prospective failure, (2) task substitution, (3) loss of timing, (4) loss of content, or (5) random error.

Scores obtained from this measure included the total overall correct PM trials, and separate subscale scores for the following: correct timed cue trials, correct event cue trials and the total correct RM score from the multiple choice questions. The total overall correct PM trials included all PM tasks (i.e., EB, TB, verbal, and action). Similarly, the
total EB and TBPM trials included both action and verbal items corresponding only to the appropriate subtype (e.g. EB or TBPM). Errors on the MISTY were analyzed as well.

Reliability analyses were conducted for both the MISTY-PM items as well as the MISTY-TBPM items. The MISTY-PM was found to have a Cronbach’s alpha value of .76 and a Spearman-Brown coefficient of .77. This suggests that the overall MISTY scale demonstrates acceptable internal consistency and split-half reliability. In contrast, the MISTY-TBPM items were found to have both a Cronbach’s alpha and Spearman-Brown coefficient value of .62, suggesting that the MISTY-TBPM scale has questionable internal consistency and split-half reliability.

**Parent Report Prospective Retrospective Memory Questionnaire for Children (PRMQC)**

This brief (16-item) questionnaire, adapted by Talbot and Kerns (2014) from the original PRMQ designed by Smith and colleagues (2000), measures the frequency of typical PM and retrospective memory (RM) failures in children’s everyday lives. In this parent-report version, parents are asked to respond to statements about their child’s memory, including “Does your child forget to either bring or turn in his/her homework that is completed?” (PM) and “Does your child forget something that he/she was told a few minutes before?” (RM). There were eight PM statements and eight RM statements with many of the PM examples encompassing both EB and TBPM (*specific* and *general*). Higher scores on the scale corresponded to higher parent-perceived frequencies of their children’s memory failures.

Talbot and Kerns (2014) demonstrated that the PRMQC has good internal consistency reliability with a Cronbach’s alpha value of 0.93 for the Total Scale, 0.91 for the Prospective scale, and 0.81 for the Retrospective Scale. Similarly, the current study
found the PRMQC to continue to show good reliability with a Cronbach’s alpha value of 0.91 for the Total Scale, 0.88 for the Prospective scale, and 0.76 for the Retrospective scale.

The data obtained from this measure were the total PM score and the total RM score. The total PM score was compared to both the overall MISTY score and the total number of completed PM tasks on the general TBPM task to determine the relationships between each of the measures.

**The Working Memory Test Battery for Children (WMTB-C)**

This published standardized measure of WM was developed by Pickering and Gathercole (2001) for children and adolescents aged 5 to 15 and was based on the 3-component structure of the *Working Memory Model* proposed by Baddeley and Hitch (1974). For the current study, participants completed two subtests (e.g., Listening Recall and Counting Recall) designed particularly to assess the central executive (supervisory system that controls information flow from and to the phonological loop and visuospatial sketchpad). During the Listening Recall subtest, participants listened to number of short sentence sets (increasing in length), and responded at the completion of each sentence as to whether it was true or false (e.g., “Cows can fly”, to which the participant should respond ‘false’). After hearing an entire set of sentences, participants were asked to recall the last word of each sentence in the trial in the order the sentences were presented. In the Counting Recall subtest, participants were shown sets of pages with a number of dots on each page, with an increasing number of pages per set. Participants were asked to count and report the number of dots pictured on each page. After counting the dots on all of the pages in the set, they were then required to report the number of dots on each page.
in the set, in the order that they were shown. Standardized age corrected scores and percentiles of each subtest were obtained as a general assessment of WM.

**General TBPM Task**

This novel task was developed as a parent-child collaborative task to measure general TBPM in a naturalistic setting. For this task, children were required to ask their parents to play the card game Go Fish sometime within a “target day”. Parents were notified of the “target day” via email 24 hours ‘prior’ to the children being provided with the request to complete the task. The general instructions for the general EPBM task were explained by the researcher during the laboratory-portion of the experiment (as outlined above). Parents and children were told that their task was to play a single game of Go Fish at any time during the target day. Twenty-four hours prior to the target day, the researcher emailed parents asking them to tell their children (as soon as they received the email in the morning) to remember to complete the task at any time between the time that they woke up to two hours after their stated bedtime on the following day. After each trial of the task, the children were asked to keep a tally of who won each of the games (i.e., them or their parent). Target days for the general TBPM trials were randomly selected to minimize the chance of habitual remembering. Thus, participants may have been asked to engage in a trial of the general TBPM task on any day at any time within normal out-of-school waking hours (i.e., before school, after school, or before bed).

Following completion of a successful general TBPM trial, parents helped their children fill out a brief questionnaire regarding that specific TBPM trial, asking how/why they remembered the task. Parents, within 2 hours after their child had gone to bed, emailed/texted the researcher to indicate their child’s completion of the trial. If the
researcher did not receive notification, a follow-up email was sent to the parent enquiring about the non-completed PM trial, inquiring about whether the parents had prompted their child to perform the task the day prior to the target date. If the parents had forgotten to deliver the cue to perform the trial, the researcher asked them to try again the following day. If the parents had provided the cue and the child had forgotten to complete the TBPM trial (e.g., initiating the game), the trial was recorded as failed and the child’s name was not entered into the raffle for that trial only. This was done in an effort to mimic a natural consequence of having not remembered a *general* TBPM task. The entire second phase of the task (naturalistic *general* TBPM task) included six to eight different target days in which six trials were completed over the course of three weeks.

The data collected from this task were the total number of successful PM trials and the contexts of children’s remembering and forgetting. The current study found the *general* TBPM measure to show poor internal consistency reliability yielding a Cronbach’s alpha value of .48.

**Completed Task Questionnaire**

Following each *general* TBPM trial, the children were asked to complete a questionnaire based on whether they had remembered to complete the task. If they had remembered the task, they were asked to fill out the Completed Task Questionnaire that consisted of one multiple-choice question that enquired as to why/how (i.e., context) they had remembered to complete the trial. There were five choices for the children to select from: (a) “mom or dad reminded me”, (b) “it just popped into my head”, (c) “I was trying to remember what I needed to do”, (d) “I saw something that helped to remind me”, or (e) “other” where they were able to add the context themselves. Coding of these responses
was completed using the following response categories by Szarras and Niedźwieńska (2011): (1) accidental rehearsal (e.g., answered $a$ or $d$), (2) self-initiated rehearsal (e.g., answered $c$), (3) no-trigger rehearsal (e.g., answered $b$), or (4) ‘other’ response (answered $e$) to aid in comparison of responses across studies.

The decision to include the specific multiple-choice questions offered in the questionnaire was based on pilot data that had been previously collected. The majority of the participants in the pilot study had chosen one of the above options to best represent the circumstances for their remembering. Additionally, there did not appear to be a lot of consistency in the minority of “other” responses provided. Similarly, most of these “other” responses could have fit within one of the other choices provided. For example, one child stated that she was going through her calendar for the day in her head and that was the reason why she remembered that she had to complete the task. Although her response provides more insight into her specific strategy for her successful prospective remembering, it still falls within Szarras and Niedźwieńska’s (2011) category of “self-initiated rehearsal”. For this reason it was felt that the response options were sufficient to encompass the majority of children’s responses.

**Forgotten Task Questionnaire**

When children forgot to complete a *general* TBPM trial, parents were asked to help their child complete a different questionnaire the following morning asking him/her to recall any circumstances of why he/she forgot to complete the TBPM trial. Unlike other studies, children provided information on their forgetting. When children forgot to complete the task, they chose either (a) “I was too busy doing other things”, (b) “I completely forgot”, (c) “Nobody reminded me”, (d) “My parent wasn’t around to play the
game with”, or (e) “Other”. Following this, children stated whether they had remembered to complete the task at any point during the day and if so, approximately how many times had they done so (i.e., 1, 2-4, 5-9, or >10).

Consistent with the Completed Task Questionnaire questions, the Forgotten Task Questionnaire questions were also pilot tested and were found to best represent the vast majority of children’s responses. Interestingly, even fewer “other” responses were endorsed on the Forgotten Task Questionnaire than on the Completed Task Questionnaire and of these, most could be viewed as fitting one of the other response options. For example, one child stated that he forgot because he spent the night at a friend’s house. Based on this information, the child could have chosen any one of the following: “I was too busy…”, “I completely forgot”, or “my parent wasn’t around…”. Therefore, the response options from the pilot study were retained for the current study.

**Procedure**

All participants individually completed the MISTY, the Wechsler Abbreviated Scale of Intelligence-II (WASI-II), and the Working Memory Test Battery for Children (WMTB-C) with the researcher present in a quiet room (either at the University of Victoria or another mutually agreed upon location). The order in which the tasks were administered was varied to minimize any potential order effects. Concurrently, their parents were asked to complete a child information questionnaire as well as the PRMQC. Following this, the parent and child were asked to complete a general TBPM task together at their own home on 6 to 8 different occasions over the course of 2.5 to 3 weeks.
During testing, a counterbalanced number of children first completed the approximate 20 minute MISTY laboratory-based PM task, following adequate understanding of the instructions. Following this, the children completed the WASI-2 and two subtests from the WMTB-C, taking between 90 and 120 minutes to complete depending on the child’s age and ability. At the end of the lab session, parents and participants were debriefed about the specifics of the experimental paradigm and the researcher demonstrated and explained the instructions of the naturalistic general TBPM task to both parent and child to ensure that they fully understood the specifics of their roles in the task (see Appendix 1 for specific details on how children were informed about the task).

Following the laboratory visit, an email/text message was sent to the designated parent’s personal mobile device/computer that contained an additional set of written instructions for the task and provided the first general TBPM target time frame (e.g., sometime on Thursday). For the first PM trial, the parent was asked to scaffold and help facilitate the task for his/her child to ensure that the child understood the task demands before performing the general TBPM trials independently. After this, the day before a general TBPM trial, the parent received an email/text message with the PM target time frame that he or she should relay to their child. During the course of the study, children engaged in the general TBPM task for six to eight days over the course of three weeks. Each PM task trial took a maximum of 30 minutes to complete and required the participation of both the child and the designated parent. Following each individual PM trial completion, the parent sent an email/text to the researcher that indicated whether the child had successfully completed the task. After all of the children had completed the
general TBPM task, the winner of the cinema tickets was contacted and sent his or her prize.
Results
As PM ability has been found to develop over the school-age years, the relationships between age and the three PM measures will be analyzed. Next, PM scores from the MISTY will be compared to scores from the general TBPM task. Next, analyses of the relationships between WM and both general TBPM and MISTY-PM will be compared. The ecological validity of the general TBPM measure will be examined by comparing it to PRMQC-PM task performance. Specific findings on the role of cue-specificity on general TBPM performance will be outlined. Finally, reports of the psychometric properties of the child-specific MISTY and it relationship to the PRMQC will be examined.

Developmental Trends
Means and standard deviations for all tasks are provided in Table 1 and means and standard deviations for all tasks by year of age are provided in Table 2. Pearson correlation coefficients between age and the three PM measures were calculated (see Table 3). Age was noted to be statistically significantly related to the MISTY PM score \( [r = .559, p < .001] \), but not to total scores on the PRMQC-PM \( [r = -.169, p > .05] \) or general TBPM \( [r = .104, p > .05] \) measures. Therefore, the effect of age was only partialled out for statistical analyses that included the MISTY-PM score.

Evaluating the MISTY
Retrospective memory for the MISTY PM tasks was also assessed using a regression analysis whereby age was entered in block one and MISTY-RM scores were entered in block two. This analysis revealed that MISTY RM was a significant predictor of MISTY PM performance even after controlling for age \( [F(2, 23) = 7.44, p < .05, \eta^2 = .057] \). Similarly, when MISTY RM scores were entered in block one and both age and
ongoing task performance were entered simultaneously in block two, both age and ongoing task performance remained significant predictors of PM performance \( F_{age} (2,23) = 7.4, p < .05, \eta^2 = .183; F_{ongoing}(2,23) = 6.69, p < .05, \eta^2 = .158 \) even after controlling for RM performance. This suggests that those who were older and those who could find more words in the word search, generally scored higher on the PM task (see Table 1 for means). Additionally, a correlational analysis of the relationship between MISTY-Timed (total score on the TBPM subscale) and MISTY-Event (total score on the EBPM subscale) yielded a significant and strong positive relationship \( r = -.54, p < .001 \).

Descriptive statistics are provided for each error type: (1) prospective failure, (2) task substitution, (3) loss of content, (4) loss of timing, and (5) random error (See Table 5). It was found that prospective failures accounted for the majority (42.86%) of the total MISTY-PM errors followed by loss of timing (22.18%), loss of content (20.68%), task substitution (12.78%), and random error (14.1%) respectively (see Figure 4).

As noted in the analysis (see Table 3) of the correlations between MISTY scores and PRMQC-PM scores, the MISTY PM performance did not significantly relate to PRMQC PM performance (PRMQC-PM) \( r = -.22, p > .05 \), even after controlling for age \( r = -.15, p > .05 \). This indicates that parents of children who performed poorly on the MISTY, regardless of age, did not rate their children as having more PM failures in daily life (PRMQC-PM) (see Figure 1).

**Correlations with the General TBPM Task**

Analysis of the correlation between general TBPM and the MISTY scores revealed that overall performance on the MISTY (Total MISTY-PM) was not significantly related to general TBPM performance (total number of successful
naturalistic TBPM tasks) \( r = -0.021, p > .05 \); see Table 4]. Similarly, performance on the MISTY-Timed cues (total score on the timed-cue subscale) was also not significantly related to *general* TBPM performance \( r = -0.009, p > .05 \). Thus, participants who performed well on the laboratory-based PM tasks (MISTY-PM and MISTY-Timed), did not necessarily do well on the naturalistic TBPM task and vice versa (see Figures 1 and 2).

**Evaluating the Relationship between TBPM and WM**

Correlational analyses were conducted to investigate the relationships between all of the cognitive measures including *general* and *specific* TBPM and WM using the average raw score across both WM tasks (see Table 3). Working memory was found to be significantly related to MISTY-PM \( r = .61, p < .01 \), MISTY-Timed \( r = .62, p > .001 \), and PRMQC-PM \( r = -0.41, p < .05 \). However, WM was not found to be significantly related to the number of completed trials on the *general* TBPM task \( r = -0.14, p > .05 \) (see Table 3). As the WMTB-C measures were significantly related to age \( r = .704, p < .001 \). Next, partial correlation analyses controlling for age were conducted. Interestingly, these analyses showed that average WM raw scores were no longer significantly related to any of the additional PM measures (see Table 4 for correlations), suggesting that WM did not appear to influence children’s overall PM ability beyond that which could be accounted for by age.

**Investigating the Role of Cue-Specificity on General TBPM Performance**

The qualitative responses given by the participants on the *general* TBPM Completed Task Questionnaires were coded as (1) accidental rehearsal, (2) self-initiated rehearsal, (3) no-trigger rehearsal, or (4) other. Descriptive analyses showed that children most commonly (69%) attributed their successful remembering to no trigger rehearsals. (See
Figure 3). This means that most children stated that their memory for action “just popped into [their] head[s]”. Additionally, when children forgot to complete their tasks, the majority (47%) stated that it was because they were “too busy with other things” (see Figure 5). Interestingly, when they did forget, only 36% said that they had remembered at some point during their day. Of those 36% who did remember during their day, 50% remembered only one time, 25% remembered 2 to 4 times, 23% remembered 5 to 9 times, and only 3% remembered 10 or more times (see Figure 6). These findings suggest that children’s everyday activities tend to interfere with their ability to remember to carry out PM tasks so much so, that children either hardly think about the task at all whilst going about their day or they choose to do more important or reinforcing things instead.
Discussion
The present study examined a subtype of TBPM, \textit{general} TBPM, in school-aged children using a novel and naturalistic PM task paradigm. Evaluation of the ecological validity of the novel \textit{general} TBPM task comparing this task to a parent-report questionnaire of everyday RM and PM failures, as well as a diagnostic laboratory-based PM measure. The study aimed to differentiate \textit{general} TBPM from \textit{specific} TBPM by demonstrating that children’s performance on the \textit{general} TBPM task was unrelated to a task that assessed \textit{specific} PM, MISTY-PM scores. Finally, this study expanded the investigation of PM in children by obtaining qualitative information about the naturalistic circumstances surrounding a child’s prospective remembering.

Results Summary
Age was found only to relate to variables associated with the MISTY task, specifically, MISTY-PM and -RM performance. Retrospective memory task performance remained a significant predictor of MISTY-PM performance even after controlling for age. Likewise, age and ongoing task performance remained predictive of MISTY-PM scores when RM performance was held constant. Consistent with previous studies, the PRMQC was found to be an internally consistent measure of children’s everyday PM errors. However, unexpectedly, PRMQC-PM scores were not related to MISTY-PM scores or to \textit{general} TBPM scores. As predicted, \textit{general} TBPM was also not related to the laboratory-based measure of PM. Additionally, after controlling for variance associated with age, children’s MISTY PM and \textit{general} TBPM performances were significantly related to WM performance.
Orthogonality of *General* TBPM

The *general* TBPM task was designed as a novel task assessing TBPM in a naturalistic way. As hypothesized, *general* TBPM performance did not significantly relate to the MISTY-PM scores suggesting that *general* TBPM, as defined and operationalized, is fundamentally different from *specific* TBPM as assessed by the MISTY. This finding provides evidence in support of the idea that *general* TBPM represents a unique PM ability unrelated to both the MISTY’s laboratory-based EB and TBPM tasks. Further support for this argument comes from the findings that *general* TBPM performance was not related to the most relevant subscale scores on the MISTY (timed cues). It can be argued that the lack of a specific time frame (e.g., “any time after 8am on Thursday…” instead of “in 2 minutes, at 10:15…”) is sufficient to differentiate *specific* TBPM performance from *general* TBPM performance. More studies focused on differentiating these proposed TBPM subtypes are required before a definitive conclusion can be reached about the orthogonality of *general* and *specific* TBPM.

In order to evaluate the second part of the argument that laboratory-based TBPM tasks are essentially EBPM tasks with the clock providing the ‘event’, an analysis comparing performance on the EB and TB MISTY tasks was conducted. Consistent with this theory, the EB and TB MISTY scores were found to be significantly related to each other. Given that previous studies have found moderate, yet insignificant relationships between EB and TBPM (Martin, Kliegel, & McDaniel, 2003; Talbot & Kerns, 2014, Yang, Chan, & Shum, 2011), it is reasonable to suggest that both PM subtypes may represent different levels of the same construct that differ in terms of the degree of self-monitoring they require. Accordingly, the EBPM task of drawing a house when one is handed a blue marker would represent a lower level EBPM task as it requires no self-monitoring to
successfully complete. In contrast, a higher level EBPM task might consist of reminding the examiner to do her homework at 10:25, 10 minutes after the task instruction had been given with a clock in open view for the participant. As this latter task requires self-initiation to disengage from the ongoing task (checking the clock to ensure that the target time had not yet elapsed), it would require more cognitive resources (self-monitoring) and thus, be more difficult to complete (Einstein & McDaniel, 1996; Ellis, 1996; Kvavilashvili & Ellis, 1996; Yang et al., 2011). However, because the clock was placed close to the participant, and did not require the participant to actively engage with it to show the time, the mere presence of the clock may have served as a trigger to remind the participant of the task. This may help to explain the large relationship observed between EB and TB performance on the MISTY. If the MISTY had required the participants to press a button to see the time, perhaps the observed relationship would have been smaller.

Currently there is a lot of support for the argument that EBPM is easier than TBPM to perform (see above). Nonetheless, the current study is the first to hypothesize that EB and TBPM both represent a continuum of truly event-based PM. This hypothesis is argued primarily because laboratory-based TBPM tasks require an event (e.g., clock being at 2:00pm) as a trigger for the memory for action. By developing a task that was less reliant on a specific time or specified brief time interval, it was reasoned that a non-event biased measure of TBPM could be obtained which might capture the way TBPM is utilized in the daily life of children. Other researchers should further investigate this idea as it has the potential to significantly alter the way PM researchers investigate and understand both EB and TBPM.
Ecological Validity of the General TBPM Task

According to Franzen and Wilhelm (1996) there are two general aspects of ecological validity: (1) verisimilitude and (2) veridicality. To evaluate the verisimilitude of the newly developed general TBPM task, we must first evaluate the extent to which the task taps into the particular behaviour domain (i.e., PM). As discussed, it is our opinion that the requirement of having to remember to carry out an action at some point during a 24-hour period is most representative of the types of PM demands people carry out on a daily basis. Likewise, as the Child Information Questionnaire responses indicated, most children and parents were familiar with card games with the majority of parents stating that they often played with their child. This information suggests that the task the participants were asked to engage in was one that they might have carried out on a normal day. Additionally, it is common for parents to ask their children to remember to do something. That said, the PM task of having to remember to ask their parent to play a game with them was unlikely to be a PM task they would be required to carry out on any given day. Similarly, children’s knowledge of being involved in an experiment might have also altered how much the task resembled a typical everyday PM task. Nevertheless, upon initial investigation, it would appear that the general TBPM employed in the current study satisfies the verisimilitude criterion necessary for establishing its ecological validity.

In order to establish the veridicality of the general TBPM task, its relationship with the PRMQC (a measure of the frequency of everyday PM and RM errors), must be determined. A significant relationship between these measures is believed to indicate that performance on the new general TBPM measure adequately reflects a person’s ability to perform daily PM tasks. Unfortunately, contrary to initial hypotheses, the current study
failed to find a significant relationship between *general* TBPM scores and daily PRMQC-PM errors. Although this appears to provide evidence against ecological validity, it is still possible that the PRMQC is not itself an ecologically valid measure of PM (see below for a discussion). Similarly, Franzen and Wilhelm (1996) have argued against the use of assessment measures that measure deficits (like the PRMQC) as criteria against which to evaluate ecological validity. They believe that researchers should use test measures that have been found to consistently relate to the behaviour and not just with deficits in order to attain the most accurate account of ecological validity. That said, Franzen and Wilhelm (1996) also state that researchers that assess the relationship between their new test measure and that of self-report, other-report, or questionnaire measures usually provide the strongest evidence of ecological validity. The lack of relationship between the PRMQC and *general* TBPM is therefore something to note. Perhaps a lack of variation due to the typicality of the children or the small sample size can better explain our findings. Additionally, many of the PM items on the PRMQC were vague and could have been interpreted as either event-based or time-based, with only a small subset best representing *general* TBPM. This may also help to explain the observed lack of significance in the relationship between PRMQC-PM and *general* TBPM.

Finally, before making a definitive conclusion about the ecological validity of our *general* TBPM task, we must consider the limitations in our evaluation (see below for a more thorough review of the limitations of the study). Firstly, this is the first study to investigate the construct of *general* TBPM. There is thus a very narrow informational basis, insufficient to rule out the possibility that the *general* TBPM task is related to everyday PM abilities. Secondly, as was mentioned earlier, PM is a complex behaviour
that is reliant on several different constructs for its successful completion (i.e., self-monitoring, WM, attention). As such, everyday PM ability is likely moderated by numerous other variables not analyzed in the current study which may have inadvertently lead researchers to over- or under-estimate a child’s actual PM ability (Franzen & Wilhelm, 1996). Finally, the “controlled” aspect of our assessment situation (e.g. independently remember to ask your mom/dad to play a game with you) did not allow children to use external devices or compensatory strategies to help them remember to carry out their tasks as that would have provided a measure of EBPM - a subtype of PM believed to be orthogonal to TBPM (Kvavilashvili & Ellis, 1996). As a result, children’s performances might better represent their internal general TBPM ability, but might not necessarily reflect their everyday PM abilities. This is because some children rely on others for reminders and others have instead developed compensatory strategies to ensure that they remember to remember (e.g. putting the cards on their bed, so that they remember in the morning).

Overall, the ecological validity of the general TBPM task cannot be properly evaluated due to reliability and validity issues associated with the measures chosen to validate it. Additionally, internal consistency of the general TBPM task was rather low, suggesting that there was little consistency in the children’s ability to remember to carry out the general TBPM task across the trials. This further suggests that other factors might better explain their overall performance on this task (e.g., motivation or other daily demands), which, in turn, might explain why the current study failed to demonstrate a significant relationship with the PRMQC. Perhaps the general TBPM task should be considered a measure of the effect of
daily life demands on general TBPM ability rather than a measure of general TBPM per se. Future research should address this possibility by investigating children’s general TBPM abilities across different environmental contexts and situations. As cognitive activities in our daily life are influenced by a multitude of external factors, it may not make sense to evaluate cognitive abilities in isolation from these factors (Goldstein, 1996). A better working knowledge of the differential effects of the environment on PM behaviours will help researchers understand everyday PM performance.

**Evaluation of the PRMQC**

Consistent with previous findings, our modified version of the original PRMQ was found to be a reliable (internally consistent) measure of PM and RM errors in children with Cronbach’s alpha values similar to those reported by Kliegel and Jäger (2007) and Talbot and Kerns (2014). The current study is the first to explicitly investigate the relationship between the PRMQC and age. Surprisingly, PRMQC scores were not found to relate to age. A possible explanation for this lack of relationship could be that the reporting parents were unconsciously comparing their children’s PM performance to same-aged peers (e.g., answering the questions based on how they felt their children performed in comparison to other children of that age, versus some absolute performance). If so, this would control for any variation in scores that would be attributable to age differences across participants.

A study conducted by Rönnlund, Mäntylä, and Nilsson (2008) used confirmatory factor analysis on the original version of the PRMQ to determine factor structure of the scale. Their findings supported a 3-factor model with PM and RM being orthogonal factors and
episodic memory acting as a common factor. Unlike their study, the current study found both the RM and PM subscales to be significantly related, with RM actually accounting for 71% of the variability in PM scores. This suggests that children’s everyday PM errors are, to a larger extent, due to an RM deficit, rather than being a product of a pure PM deficit. According to Ellis (1996), RM ability should only account for a small proportion of a person’s ability to successfully carry out a previously formed intention as PM is a complex construct relying on multiple factors for its successful completion (see introduction for a review of this process and its relationship to other cognitive abilities) which suggests that the PRMQC may be assessing slightly different abilities than the adult measure. That said, if a person has a RM deficit such that he or she is unable to properly encode the PM task, it would not matter if all of his or her cognitive abilities are properly functioning. This person would experience a prospective failure in that he or she would be unable to recall the content of the intention. Because the RM scale is strongly related to the PM scale, it could be argued that the children’s PM difficulties, as reported on the PRMQC, were actually a result of encoding difficulties. This argument is further supported by the MISTY error analysis that demonstrated that prospective failures (e.g., having a PM task completely slip one’s mind) accounted for the majority of the errors children made when performing the MISTY task.

The lack of relationship between the MISTY PM measure and PRMQC warrants further discussion. This finding is in direct contrast with the majority of studies investigating PM using the original PRMQ. These studies have consistently found the PRM-Q-PM scale to significantly correlate with laboratory-based PM measures (Kliegel & Jäger, 2006; 2007; Mäntylä, 2003; Talbot & Kerns, 2014). That said, some of these
studies have found the two subscales (RM and PM) to be differentially sensitive in predicting actual PM performance (Kliegel & Jäger, 2006), whereas others have not (Kliegel & Jäger, 2007; Mäntylä, 2003). Kliegel and Jäger (2006) further discovered that the differential sensitivity of their PRMQC scales was related to the subtype of PM that was being measured (i.e., EB or TBPM). Interestingly, the other studies that failed to demonstrate differential sensitivity predicting their laboratory-based PM tasks only included one subtype of PM, either EB or TBPM. As the MISTY-PM score combined both EBPM and specific TBPM task scores, perhaps the observed lack of differential sensitivity in PRMQC PM and RM subscales was due to a lack of variation in PM subtypes. It is possible that analyses using separate MISTY scores for EBPM and TBPM tasks respectively might yield differential sensitivities across the two PRMQC subscales.

**Evaluation of the MISTY**

The MISTY, like its precursor the MIST, was designed to identify children and youth with substantial PM deficits. For this reason, the lack of relationship between this measure and the PRMQC could potentially be explained by the characteristics of the population studied. A criticism of the use of diagnostic neuropsychological measures is that they tend to be designed as diagnostic indicators and often yield unreliable measurements of cognitive abilities in unimpaired or superior ranges of performance (Franzen & Wilhelm, 1996). Because participants in the current study specifically included only those who were typically developing, they were not likely to exhibit the range of variability in PM that might be seen in children with deficits in this area. As such, it is probable that due to low variation in the MISTY PM scores the likelihood of finding a significant correlation with PRMQC scores was reduced.
**Relationship between General TBPM and WM**

Previous studies have consistently shown WM to significantly predict laboratory-based TBPM performance (Kerns, 2000; Mackinlay et al., 2009; Mahy & Moses, 2011). This makes sense from a logical standpoint as oftentimes these PM measures require participants to remember multiple different tasks in a short time frame, thus preventing sufficient time for the tasks to be fully consolidated into long-term memory. That said, studies that have systematically reduced the laboratory-based PM task’s WM load by embedding a task list within their paradigm or by allowing participants to write down the task instructions have demonstrated weaker relationships between WM and PM (Rose et al., 2010).

In the current study, The MISTY required participants to remember to spontaneously recall and perform eight different PM tasks over only 20 minutes, with delays of either 2 or 10 minutes. As memory consolidation into long-term memory can take hours to occur (Meeter & Murre, 2004), it was believed that the participants in the current study would have to rely heavily on their WM ability to successfully complete their PM tasks. Contrary to initial predictions, after controlling for age, MISTY PM scores did not significantly relate to the average WM raw scores.

Unlike the MISTY, general TBPM was hypothesized to rely less on WM for its successful completion. This was believed to be because the TBPM task instructions were given 24 hours prior to needing to be recalled, thus allowing sufficient time to fully consolidate the intended action into long-term memory. Consistent with this hypothesis, general TBPM was not found to be a significant predictor of WM performance.
Reasons for Remembering and Forgetting General TBPM Tasks

The current study found that the majority of children who remembered to carry out their PM tasks, did so because it “popped” into their heads. This finding provides support for Kvavilashvili and Fisher’s (2007) theory that TBPM retrieval can occur spontaneously, periodically popping into one’s mind, during the delay interval. The researchers argue that this “popping” is a result of a sustained subthreshold activation of its representation.

In contrast, however, the current finding does not support those found by Szarras and Niedźwieńska (2011). In particular, their study found accidental internal or external triggering to best explain their adult population’s successful remembering. Although the difference in the findings of the current study and those of Szarras and Niedźwieńska (2011) could reflect a developmental shifting in memory strategy use, it may also reflect limitations in children’s metacognitive abilities (see limitation section for a more in depth discussion). It is possible that children’s memories are actually being triggered by subtle internal or external stimuli (see the “random walk” model of PM developed by Wilkins, 1978), but their developing self-awareness limits their ability to identify the stimulus that actually triggered their memory (e.g. glancing at the television and seeing kids playing cards) (see Sternberg (2009) for a discussion of the development of self-reflection). Nevertheless, this is an interesting finding as it suggests that children’s memories are being triggered by more internal cues. This therefore has direct implications for the feasibility of developing effective interventions and/or rehabilitation strategies for children with PM difficulties.

That said, a further investigation into children’s perceived reasons for forgetting might be more helpful when considering ways in which to help children remember to complete
future tasks. As was seen in our findings, the children in our sample tended to find themselves distracted with their daily activities to the point that they hindered their ability to even think about the PM task. Taken together with the findings on the context for children’s successful remembering, one might suggest that children’s difficulties completing PM tasks stem more from difficulties with task switching – an executive function that has previously been found to influence PM in adults (Kliegel, McDaniel, & Martin, 2003; West, Scolaro, Bailey, 2011). Nonetheless, the current study is the first of its kind to ask children to retrospectively report on the contexts for their remembering. For this reason, caution should be used when interpreting our findings. Further studies employing such methods is advised as the information gleaned has the potential to be very useful to researchers who are looking to better understand people’s internal PM experiences. From there, they may develop more effective interventions and/or rehabilitation strategies for children with PM problems.

**Limitations and Future Research**

Given that this study was one of the first of its kind, there were several limitations associated with it. These limitations can ultimately be subdivided into two separate categories, namely participant-associated and paradigm-associated limitations, and are discussed in turn. In terms of participant-associated limitations, one of the most important limitations is the sample size. This issue likely limited the power of several of our analyses that, in turn, lowered the chances of obtaining significant effects.

A second limitation is the demographics of the sample. All of the participants included in the study were from the Victoria and Lower Mainland area and achieved an average full-scale IQ on the WASI-II of 112.52 (See Table 1). This suggests that the families who
volunteered their time were, on average, well-educated with an affinity for academic research. Similarly, given the long-term nature of the naturalistic paradigm and the initial time commitment associated with the cognitive testing, it is not unlikely that the parents who were initially interested in participating in the study also tended to spend more of their spare time engaged in social games with their children.

A final participant-associated limitation in the study was that a substantial number of the participant pool consisted of sibling dyads. Although protocols were put in place to ensure that the sibling(s) did not unduly influence the currently-participating child’s ability to remember the naturalistic task (i.e., engaged in naturalistic task one at a time [randomly assigned] and was explicitly asked not to help his or her sibling remember), it is still possible that prolonged exposure to the paradigm inadvertently resulted in the task becoming habitual for the second and sibling. That said, the data do not initially seem to suggest that second-participating siblings performed better than their initially-participating sibling on the naturalistic PM. Despite this, the presence of siblings in the study, does effectively limit the variability in the sample. Overall, the identified participant-associated limitations in the study work together to limit the generalizability of the findings of this study.

In addition to the aforementioned participant-associated limitations, this study is also associated with a number of limitations based on the paradigm chosen. As is inherent in most naturalistic task paradigms, the current study suffers from a lack of control over some important variables which likely impacted the reliability of the task. For example, the naturalistic task is designed to be delivered to the participating child by his or her parent as soon as the parent receives the target email from the researcher (ex., between
7:30am and 8:30am, 24 hours before the target day). However, over the course of the study, it became increasingly apparent that the time in which the parent delivered the message to the child varied considerably (ex., anywhere from 24 hours to 10 hours before the target day). Parent feedback indicated that their delay in notification often occurred because they were either already at work when they read their messages or the child was already at school. Therefore, some children would receive the message when their parent returned from work or when they returned home from school or an extracurricular activity. Ultimately, this resulted in children receiving more or less of an advantage for remembering the task, with the children who received less notice potentially relying more on their WM to remember the task than the others. However, as no data was obtained on the actual time a participant received the message, it is impossible to say whether this did influence the findings of the study. Regardless, the low reliability of the naturalistic task will limit the ability to find correlations with other measures of general TBPM, such as parent reports.

Another limitation related to the logistics of the 24-hour delay period between receiving the message and carrying out the task is that the chosen target days could not be back-to-back (i.e. Tuesday and Wednesday of one week). This is because having the parent deliver the message to his or her child on a “target day” would have served as a reminder for the child to complete the PM task for that day. Consequently, the majority of the seven target days occurred every other day with the exception of one or two tasks that consisted of a two-day delay. Although the target days were chosen in an effort to minimize the likelihood of habitual remembering, the limited variation in the target day schedule was constrained by the three week maximum duration of the task and as such,
may have been unduly influenced by habituation after all. Similarly, some of the participants in the study came from separated or divorced families with varying custody and visitation arrangements. This means that some participants’ PM task schedules were dependent on the days the child was in the participating parent’s custody, further limiting the researcher’s ability to standardize the PM task schedule across participants.

Task motivation was another factor that may have served to limit the findings of the study. As mentioned, the card game of Go Fish was chosen as it was a game that the researcher felt was neutrally motivating and that children across the age range of the study would be approximately equally familiar with. However, throughout the study, it was evident that some children were not at all interested in playing Go Fish and asked to play a different game instead. In these instances, the researcher instructed the child to choose a different game that was not highly motivating or rewarding so as to not bias the child’s memory. Although the researcher instructed the child to do this, there was no way of controlling or assessing children’s intrinsic motivation to play the game. Therefore, differences in intrinsic motivation may have at least partially contributed to the likelihood of remembering the PM task.

In addition to intrinsic task motivation, external task motivation may also have influenced the findings of the study. More specifically, the children’s motivation to play a card game in general may not have been sufficient enough to warrant them remembering to conduct the task. For example, children did not suffer any real life consequence for not remembering to play the game with their parent apart from not earning an entry into a raffle for this their trial. Furthermore, the children in the study may not have had a sufficient understanding of what a “raffle” truly is and what forgetting a trial actually
meant for their chances of winning the pair or movie tickets. Therefore, children may not have been sufficiently externally motivated to conduct the general TBPM task to the best of their ability and as such, their scores may not provide the most accurate measurement of their true real-life general TBPM ability.

A final limitation of the study is associated with the information gleaned from the qualitative forgotten task and task completion questionnaires. An important limitation of the use of these questionnaires was the age range of the participants who filled them out (e.g. 7 years, 1 month, to 13 years 5 months). Within this age range, there are a number of metacognitive skills (i.e. self-reflection) that are developing (Sternberg, 2009). This may have contributed to differences in the way the children responded with older children having an increased ability to more accurately reflect their experience than younger children. Additionally, as soon as the children completed the task, they filled out the Completed Task Questionnaire, but when they forgot a task, they would often respond to the Forgotten Task Questionnaire the next day. Although parents were good about having their children fill out the forgotten questionnaire as soon as possible, the increased delay in responding may have introduced some retrospective memory failure. However, given that the questionnaires were in a multiple-choice format, the response choices may have sparked their recognition memory for the task and counteracted any potential recall failures. Finally, I had to quantify the children’s qualitative questionnaire responses in order to statistically analyze them. As a consequence, the information that could be gleaned from children’s responses was limited, which also limited the depth of my understanding of their subjective PM experiences.
In summary, the naturalistic paradigm chosen was the first of its kind and as such, did suffer from a number of limitations that future researchers should address. Despite this, the findings suggest that these types of tasks provide additional information about why and how children successfully remember to complete their daily PM tasks. By incorporating more of these types of paradigms, future research may be better able to identify areas for PM intervention in children who appear to struggle in this area.
References


Kliegel, M., Brandenberger, M., Aberle, I. (2010) Effect of motivational incentives on


Mills, G., Garbarino, J., & Raskin, S. (personal communication). Measurement of
prospective memory in children: The memory for intentions screening test for youth (MISTY).


Appendix 1

Description of How the General TBPM Task Was Explained

Following completion of the children’s cognitive testing, children were rejoined with their parent and the researcher outlined the details of the at-home task to both of them. First, the researcher explained what the child would be expected to remember to do (e.g., engage their parent in a game of Go Fish at any point on a “target day”). The researcher then described what was meant by a “target day” (e.g., 24 hours after the parent received an email from the researcher). To ensure understanding, the researcher provided a more concrete example (e.g., “If I send your mom an email between 7:30 and 8:30am tomorrow morning [Monday], you will have to remember to ask her to play Go Fish with you anytime on Tuesday”). The researcher further explained that the child would not be expected to carry out the task while at school/camp or on any day that he or she was away from his or her parent. Additionally, the researcher explicitly stated to the child that they were to try and remember all by themselves without the use of any external reminders (e.g., parents, siblings, friends, notes, alarms, etc.)- explaining that for the research it was important to know how well the child could remember using only his/her brain.

To ensure that the child sufficiently comprehended task expectations, the parent was asked to complete the first task as a real-life example. The researcher provided another concrete example (e.g., “So as the practice and real life example, if I send you [the parent] an email tomorrow [Monday], you [the parent] will be responsible for remembering to engage your child in the game at any point the next day [Tuesday] so that he/she knows what he/she is expected to do for the next 6 tasks”).

If during the description of the task, either child or parent communicated disinterest in playing the card game chosen for the study (e.g., Go Fish), the researcher ensured them
that they could choose a different game. However, she also emphasized that the game that
the child decided to choose had to be of neutral motivation, given that the objective of the
game was not to be “too exciting or motivating” or “completely boring” for the child. The
researcher explained that the reason for this was to ensure that the game choice did not
accidentally increase their likelihood (or unlikelihood) of remembering to complete the
task.

After the researcher had ensured that both the child and the parent had adequately
understood the task procedures, the questionnaires were then explained using a blank
copy of the response form. The researcher stated that once the child had completed the
task, he/she would then have to state why/how he or she remembered to complete the
task. The parents were then asked to email the researcher once this was completed to
ensure that the researcher could keep track of performance. Parents were then notified
that the researcher would send a reminder email if she had not received notification from
the parent of his or her child’s task completion. At this time, the researcher explained that
the child would earn an entry in a raffle for a grand prize for every successfully
completed task.

The researcher also outlined what the child and parent should do if the child
accidentally forgot to complete the task. The researcher stressed that forgetting was likely
and that the child should not feel upset or concerned if they forgot to complete a task. She
also emphasized that learning why children might forget was a part of the experiment,
and then showed and explained the Forgotten Task Questionnaire. After this was
sufficiently explained and the researcher felt the child and parent to have a good grasp of
the task, it was once again emphasized that the consequence of not remembering to complete the task was that the child did not receive a raffle entry for that specific task.

The researcher then queried for any final questions and addressed these. The parents were then requested to help ensure that neither the questionnaires nor the child’s game of choice would be placed/kept in the child’s immediate view during the study, as this would engage a different type of prospective memory. Additionally, parents were asked to ensure that the child’s siblings and/or other parent would not help to remind the child to complete his or her tasks. The researcher then asked the child to try his or her best to remember as many as possible, relying only on their own remembering. Following this, the researcher sent the parent an email that outlined the task instructions in point form so that the parent could refer to it when needed.