

The delay period as an opportunity to think about future intentions: Effects of delay length and delay task difficulty on young adult's prospective memory performance

Caitlin E. V. Mahy¹ · Katharina Schnitzspahn² · Alexandra Hering³ ·
Jacqueline Pagobo¹ · Matthias Kliegel³

Received: 1 September 2016 / Accepted: 4 January 2017 / Published online: 2 February 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract The current study examined the impact of length and difficulty of the delay task on young adult's event-based prospective memory (PM). Participants engaged in either a short (2.5 min) or a long (15 min) delay that was filled with either a simple item categorization task or a difficult cognitive task. They also completed a questionnaire on whether they thought about the PM intention during the delay period and how often they thought about it. Results revealed that participants' PM was better after a difficult delay task compared to an easy delay task. Participants thought about the PM intention more often during the difficult delay task than during the easy delay task. PM performance was positively related to participants' reports of how many times they thought about their intentions. The important role of delay task difficulty in allowing or preventing individuals from refreshing their future intentions is discussed.

Introduction

The delay period has an opportunity to think about future intentions: Effects of delay length and delay task difficulty on young adult's prospective memory performance.

In daily life, individuals often have to remember to carry out their future intentions such as mailing a letter the next time they see a mailbox or dropping off a gift at a

friend's house after work. This ability, known as prospective memory (PM; Einstein and McDaniel 1990), is critical to successful occupational and social functioning. Two PM task types have been described: time-based PM tasks that rely on an intention being carried out at a particular point in time or after a specific amount of time has passed and event-based PM tasks that rely on intentions being carried out in response to a specific cue in the environment.

The challenge in carrying out PM tasks comes from the fact that such tasks do not exist in a vacuum but are delayed in nature and thus distracting tasks must be completed in the interim before the appearance of the PM cue. Dropping off a gift at a friend's house right after work might be considered a relatively simple task after a quiet afternoon at the office, but this task is normally seen as much more difficult after a busy afternoon filled with meetings, returning phone calls, answering emails, and picking up your children on the way home. While these assumptions seem plausible, surprisingly, empirical research has revealed mixed results for the role of the delay interval on PM performance. Thus, the current study will further examine the impact of two key components of the delay interval, its length and difficulty level of the intervening activity, on event-based PM performance in young adults directly testing intention refreshing as a potential supporting mechanism.

Length of delay interval

The impact of the length of the delay interval on event-based PM has been examined by several researchers, but has revealed contradictory findings (see Martin et al. 2011 for an overview). Although the majority of studies have found no effect (e.g., Einstein et al. 1992; Guynn et al. 1998; Stone et al. 2001) or a negative impact of a longer delay interval on PM performance (e.g., Brandimonte and

✉ Caitlin E. V. Mahy
caitlin.mahy@brocku.ca

¹ Department of Psychology, Brock University, 1812 Sir Isaac Brock Way, St Catharines, ON L2S 3A1, USA

² University of Aberdeen, Aberdeen, UK

³ University of Geneva, Geneva, Switzerland

Passolunghi 1994; Johansson et al. 2000; McBride et al. 2013; Nigro et al. 2002; Scullin and McDaniel 2010; Somerville et al. 1983; Tierney et al. 2016), two studies have found a positive impact of a longer delay on PM (Hicks et al. 2000; Mahy and Moses 2011).

Hicks, Marsh, and Russell (2000) examined the impact of delay interval length on young adults' PM in a series of experiments. In the first experiments, participants in the short delay condition were asked to read and rate humorous cartoons for 2.5 min, whereas participants in the long delay condition completed the remote associates test and the Raven's Progressive Matrices, a measure of fluid intelligence (Raven 1941) for 15 min. Results showed that PM performance was *better* after a long delay compared to a short delay. A third experiment examined delay intervals of 2.5, 5, and 15 min, and had participants complete a vocabulary measure where they had to choose synonyms and antonyms. Results of this experiment again showed a positive effect of longer delay interval length on PM performance. Similarly, Mahy and Moses (2011) have replicated this positive impact of delay length on PM performance in preschool-aged children. Findings from their study revealed that 5 years old performed better after a 5-min delay compared to a 1-min delay, whereas 4 years old were not impacted by the delay period. These results were interpreted to show that 5 years old might have the abilities necessary to take advantage of the delay period by monitoring or refreshing their PM intentions, whereas 4 years old might have not yet developed this ability.

Although theories of memory decay in the classic retrospective memory literature would predict worse PM performance after a longer delay interval (e.g., Brown 1958; Ebbinghaus 1885, 1964), Hicks, Marsh, and Russell (2000) suggested that the increases in PM performance with longer delays might be due to individuals having more time to refresh or reflect on their PM intentions. Notably, they argue that easier delay tasks should provide more opportunities to think about one's PM intention compared to more difficult delay tasks that may not afford as many opportunities to refresh one's future intentions. Refreshing one's intentions during the delay period should boost cue generation and strengthen cue–response associations due to increased accessibility of the cue leading to better PM performance (see Souza et al. 2014 for a similar argument for refreshing material in visual working memory). Thus, the extent to which one can refresh their prospective intentions during the delay might have a more important influence on PM performance than delay length.

Several researchers have found that varying delay lengths has no impact on PM performance in adults (e.g., Einstein et al. 1992; Guynn et al. 1998; Stone et al. 2001). In a sample of younger and older adults, Einstein et al. (1992) had participants complete several retrospective

memory tasks during a delay or complete the same tasks plus an additional 15-min task. Findings revealed that subsequent PM performance did not differ across these two conditions despite the 15-min difference in delay interval length between them. Importantly, in this study, the longer delay interval had additional content that the short delay interval did not contain. Similarly, Guynn et al. (1998) found in a series of four studies that the delay between the PM instructions and the PM task or between a PM reminder and a PM cue appearance (ranging from 1 to 20 min) did not impact young adults' PM performance. Notably, for many of the experiments, the content of the short and long delay conditions were not matched.

Finally, some naturalistic studies have shown a negative impact of delay on PM. For example, Johansson, Andersson, and Ronnberg (2000) found that older adults had worse PM after a long delay (60 min) compared to a short delay (10 min). Participants had to remember to remind the experimenter what to do when passing a certain location on campus or indicating when a certain amount of time had elapsed. Similarly, a naturalistic study with 2- to 4-year-old children found that children were better at reminding their mothers to do something after shorter delays (5–10 min) compared to longer delays (6–12 h; Somerville et al. 1983). Importantly, in everyday life, individuals are more likely to complete several different tasks during a longer delay compared to a shorter delay. Thus, the effect of the number of tasks and complexity of tasks completed during the delay is impossible to control for in such naturalistic studies.

In sum, there is mixed evidence on the impact of delay length on PM performance that seems to differ depending on the nature of the PM task and the activity that fills the delay interval. Importantly, the majority of the studies that have examined the impact of delay length on PM performance have confounded delay length and delay task difficulty without taking an independent measure of task difficulty. Thus, it is challenging to pinpoint what is driving these delay effects: the length of the delay or the nature of the delay activity.

Content of the delay interval

Beyond the length of the delay interval, the content of this delay interval might be an important factor to consider (see Mahy et al. 2014 for review). Similar to the research on the impact of delay length on PM performance, the findings on the effect of delay task difficulty on PM performance are mixed. Some results suggest that a cognitively demanding delay task has a more negative impact on PM performance compared to less demanding task (see Brandimonte and Passolunghi 1994; Mahy and Moses 2015) and other studies have found no effect of delay task difficulty on later PM performance (Cook et al. 2014; Shelton et al. 2013).

In a sample of adults, Brandimonte and Passolunghi (1994) found that PM performance was not affected by a delay when it was unfilled or filled with simple counting compared to when the delay interval was filled with a short-term memory task or a motor task where delay negatively affected PM. Mahy and Moses (2015) found that 4- and 5-year-old children did worse on a PM task after the delay was filled with a challenging version of a visual working memory task (self-ordered pointing task) compared to when the delay was filled with an easier version of the same task. In contrast, Shelton and colleagues found that young adults PM performance was unaffected by a difficult relative to an easy delay task, but that old-old adults showed a negative impact of delay task difficulty on PM performance (Shelton et al. 2011, 2013). Similarly, Cook, Ball, and Brewer (2014) found no impact of executive control depletion from a Stroop task on PM performance. Thus, another potentially important aspect of a delay interval is the difficulty of the intervening task, since it is possible that delay task difficulty could impact later PM performance.

One potential reason why the difficulty of the delay task as well as the length of the delay period may have important implications for later PM is that all these factors may provide opportunities for refreshing intentions as suggested by Hicks, Marsh, and Russell (2000). Depending on the nature of the delay task, individuals may be able refresh their intentions and engage in mind wandering, which might promote thinking about their PM intention. Further, very few experimental studies have followed-up with participants by asking them questions about what they did during the delay period in terms of whether and how often they thought about the PM intention which overlooks potentially important and interesting responses (although see Kvavilashvili and Fisher 2007; Szarraz and; Niedźwieńska 2011; who have done so in naturalistic settings and with self-generated PM tasks).

The current study

Despite the contradictory findings on the effect of delay length on PM, little work has systematically examined the effect of the length of delay on young adult's PM, since Hicks, Marsh, and Russell in 2000. Further, only a handful of studies have examined the impact of delay task difficulty, a potentially critical factor in allowing or limiting the extent to which individuals refresh their intentions. Moreover, conflicting results exist on the influence of delay task difficulty on PM performance. The current study sought to fill these gaps in the literature by attempting to replicate the findings of Hicks, Marsh, and Russell (2000) but also to consider another important factor—delay task difficulty—that could have an impact on whether individuals can take advantage of a delay to think about one's intentions or not.

The current study improved on previous study designs in that we measured performance on the delay task to objectively access difficulty level and maintained task difficulty across short and long delays. Finally, we sought to expand on past research by asking participants to provide self-reports on whether and how much they thought about the PM intention during the delay period and whether this intention was always present in their mind throughout the delay interval. Given the findings of the literature so far, there are three possible outcomes: (1) that a longer delay will have a negative impact on PM performance, (2) that a longer delay will have no impact on PM performance, or (3) that a longer delay will have a positive impact on PM performance. Another goal of the current study was to explore the impact of delay task difficulty on adult's PM. Here, we expected that a difficult delay task would have a negative or neutral effect on PM performance. Finally, we expected that participants' self-reports on whether and how often they thought about the intention during the delay period will positively relate to PM performance above and beyond their interest in the tasks or tiredness given the previous literature (Kvavilashvili and Fisher 2007).

Method

Participants

Participants were 140 Brock University undergraduate students (97 women; $M_{\text{Age}} = 20.18$, $SD_{\text{Age}} = 4.28$) who received partial credit toward a course requirement for their participation. Eight participants did not produce useable data due to: not being able to answer the retrospective memory question for the PM task instructions ($N=3$), misunderstanding the task rules ($N=2$), failure to pay attention to the instructions ($N=1$), experimenter error ($N=1$), and technical difficulties ($N=1$). The final sample was composed of 132 participants. Participants were randomly assigned to one of the four experimental conditions that resulted from fully crossing (between-subjects) the two levels of delay length (2.5 vs. 15 min) with the two levels of delay task difficulty (Difficult: Raven's Matrices vs. Easy: simple item categorization task). There were equal sample sizes in each condition ($N=33$). Participants were mostly Caucasian (71%), from middle-class backgrounds, spoke English fluently, had no psychological, neuropsychological problems, and had normal or corrected vision.

Procedure and materials

Participants were tested individually in a quiet testing room at Brock University. PsychoPy (version 1.82.01; Peirce 2007) was used to present the experiment to participants on

a 13-inch MacBook Pro laptop computer. Participants were randomly assigned to one of four experimental conditions that either had a short (2.5 min) or long delay (15 min) and that was filled with an easy delay task (item categorization) or difficult delay task (Raven's Matrices).

First, participants read the instructions for the ongoing lexical decision task from the computer screen. Then, participants completed five practice trials of the lexical decision task. After this lexical decision task practice, participants read about the PM instruction to press the '9' when an animal name appeared in the lexical decision task. Then, participants were asked to turn to the experimenter and explain what they had to do aloud (to confirm understanding of the ongoing and PM tasks). Once the experimenter had confirmed the participants understanding of the ongoing task and PM task rules, the delay period began. Participants either had a 2.5- or 15-min delay filled with either the easy or difficult delay task.

In the easy delay task, participants were presented with 50 photographs of plants and 50 photographs of household items on the computer screen, and were asked to judge whether an item was living or non-living. Participants were asked to press '1' when a photograph depicted a non-living thing and '2' when the photograph depicted a living thing. The 100 photographs appeared randomly without repetition, unless participants completed 100 in which case they were repeated in a randomly selected order. Performance on item categorization was self-paced and accuracy and response times were measured. Participants were instructed to complete the task as quickly and accurately as possible.

In the difficult delay task, participants were presented with a computerized version of the matrices from Raven's Progressive Matrices (Raven 1941). The 60 matrices appeared randomly without repetition, unless participants completed all 60 in which case the matrices were repeated in a randomly selected order. Participants were asked to select the picture that completed the matrix by selecting the appropriate number on the keyboard. Matrices performance was self-paced and accuracy and response time were measured. As with the easy delay task, participants were instructed to complete the task as quickly and accurately as possible.

After the delay interval, participants began the ongoing lexical decision task in which the PM cues were embedded. The prospective instruction to press the 9 key when they saw an animal word was not mentioned, but a brief instruction for the lexical decision task was provided again. On each lexical decision trial, a fixation point appeared for 500 ms followed by a word or pseudoword that remained on the screen until the participant responded. One hundred and four words/pseudowords were randomly selected from the English Lexicon Project (Balota et al. 2007). Four of these words represented animal names (i.e., penguin,

lizard, chimpanzee, and goose) and occurred on trials numbered 25, 50, 75, and 100. All words and pseudowords were presented in a fixed order including the animal words. Participants made word/pseudoword judgments of the stimuli by pressing '1' for a word and '2' for a non-word. An equal number of words and pseudowords were presented. For the PM task, participants were asked to press the '9' key instead of making a lexical decision for animal words. Accuracy and reaction times for ongoing task items and PM targets were recorded as the dependent variables. PM responses were counted only if they occurred immediately after the appearance of the PM cue.

After the PM task was completed, the experimenter probed the participant on their understanding of the PM task. First participants were asked to describe what they had to do during the lexical decision task. If participants did not mention pressing the '9' key when they saw an animal word, the following questions were asked in a fixed order until the participant could answer: (1) "What were you supposed to do in the lexical decision task?", (2) "Was there something else you had to do in the lexical decision task?", and finally (3) "What were you supposed to do when you saw an animal word?" If participants could not answer any of these probes correctly, they were excluded from the final sample. Finally, participants were given a questionnaire that included: a confirmation for their memory of the PM instruction, questions about if and how often they thought about the PM intention during the delay task, if the intention was always present in their mind, their tiredness, interest in the delay and lexical decision tasks, and some basic demographic information (see Appendix). The research ethics board at Brock University approved all procedures.

Results

Table 1 shows means and standard deviations for accuracy and reaction times for the Delay task, the Lexical Decision task (ongoing task), and the PM task by experimental condition.

Delay task performance

Overall, participants performed above chance on the simple categorization task as well as the Raven's Matrices, $t(66) > 7.39$, $p_s < 0.001$. Performance accuracy on the simple categorization task ($M=0.94$, $SD=0.16$) was superior to performance on the Raven's Matrices ($M=0.64$, $SD=0.15$; $t(130)=11.12$, $p < .001$, Cohen's $d=1.93$), confirming that the Matrices were significantly more difficult than the item categorization task. Further, participants average reaction time was significantly faster for the item categorization task ($M=726.79$, $SD=147.88$) compared to

Table 1 Means and standard deviations for performance on the delay task, lexical decision task, and prospective memory task by condition

| | Experimental condition | | | |
|---------------------------------------|---|---|--|--|
| | Condition 1 Short delay, easy delay task | Condition 2 Short delay, hard delay task | Condition 3 Long delay easy delay task | Condition 4 Long delay, hard delay task |
| Delay task accuracy | 0.91 (0.23) | 0.61 (0.18) | 0.96 (0.02) | 0.67 (0.12) |
| Delay task reaction time | 758.81 (164.38) | 7253.29 (2238.41) | 695.67 (124.46) | 7517.71 (3853.43) |
| Number of delay task trials completed | 198.24 (32.86) | 18.18 (5.24) | 1311.76 (231.03) | 122.12 (62.64) |
| Lexical decision task accuracy | 0.80 (0.08) | 0.79 (0.06) | 0.82 (0.06) | 0.79 (0.08) |
| Lexical decision task reaction time | 1321.21 (455.31) | 1402.09 (320.88) | 1420.98 (422.45) | 1485.72 (456.29) |
| Prospective memory accuracy | 0.15 (0.30) | 0.42 (0.37) | 0.23 (0.36) | 0.45 (0.37) |
| Prospective memory reaction time | 1345.54 (332.91) | 1175.31 (461.49) | 1093.29 (306.35) | 1353.30 (670.28) |

There were 33 participants in each condition. Reaction times were calculated for accurate trials only

the Raven's Matrices ($M=7385.50$, $SD=3131.06$) for correct trials only, $t(137)=17.65$, $p<.001$, Cohen's $d=3.00$, confirming that the item categorization trials were quicker to complete than the Raven's Matrices trials. Participants completed significantly more trials of the delay task in the long delay conditions ($M=716.94$, $SD=622.46$) compared to the short delay conditions ($M=108.21$, $SD=93.68$), $t(130)=7.86$, $p<.001$, Cohen's $d=1.37$; however, this difference did not have an impact on their accuracy on the delay task during short ($M=0.76$, $SD=0.25$) or long ($M=0.82$, $SD=0.17$) delay periods, $t(130)=1.49$, $p=.14$, Cohen's $d=0.28$, or on the reaction times on the delay task after a short ($M=4053.11$, $SD=3635.34$) or long ($M=4106.69$, $SD=4373.58$) delay, $t(137)=0.078$, $p=.938$, Cohen's $d=0.01$.

Ongoing task performance: Lexical decision task

Participants performed above chance on the lexical decision task, $t(131)=47.60$, $p<.001$, with high accuracy on average across experimental conditions ($M=0.80$, $SD=0.07$). For correct lexical decisions, participants responded in less than 1500 ms on average ($M=1413.73$, $SD=375.71$).

A 2 (Delay length: short vs. long) by 2 (Delay Task Difficulty: hard vs. easy) ANOVA on lexical decision task accuracy revealed no main effect of delay, $F(1, 128)=1.46$, $MSE=0.005$, $p=.23$, $\eta_p^2=0.01$, no main effect of delay task difficulty, $F(1, 128)=2.55$, $MSE=0.005$, $p=.11$, $\eta_p^2=0.02$, or interaction between delay length and delay difficulty, $F(1, 128)=0.60$, $MSE=0.005$, $p=.44$, $\eta_p^2=0.005$. Similarly, a 2 (Delay length: short vs. long) by 2 (Delay Task Difficulty: hard vs. easy) ANOVA on lexical decision task reaction time (for accurate trials only) did not reveal a main effect of delay length, $F(1, 128)=1.72$, $MSE=138959.18$, $p=.19$, $\eta_p^2=0.01$, a main effect of delay task difficulty, $F(1, 128)=3.06$, $MSE=138959.18$,

$p=.08$, $\eta_p^2=0.02$, or an interaction between delay length and delay task difficulty, $F(1, 128)=0.29$, $MSE=138959.18$, $p=.59$, $\eta_p^2=0.002$.

Prospective memory performance

Figure 1 shows performance on the PM task by condition. Performance on the PM task was quite low with overall accuracy levels just above 30% ($M=0.31$, $SD=0.37$). A 2 (Delay length: short vs. long) by 2 (Delay Task Difficulty: hard vs. easy) ANOVA on PM accuracy revealed a significant effect of delay task difficulty, $F(1, 128)=15.52$, $MSE=0.13$, $p<.001$, $\eta_p^2=0.108$, with participants performing better after a difficult delay task ($M=0.44$, $SD=0.37$) compared to an easy delay task ($M=0.19$, $SD=0.33$). There was no significant effect of delay length on PM performance, $F(1, 128)=0.97$, $MSE=0.13$, $p=.327$, $\eta_p^2=0.008$, and no interaction between delay length and delay task difficulty, $F(1, 128)=0.14$,

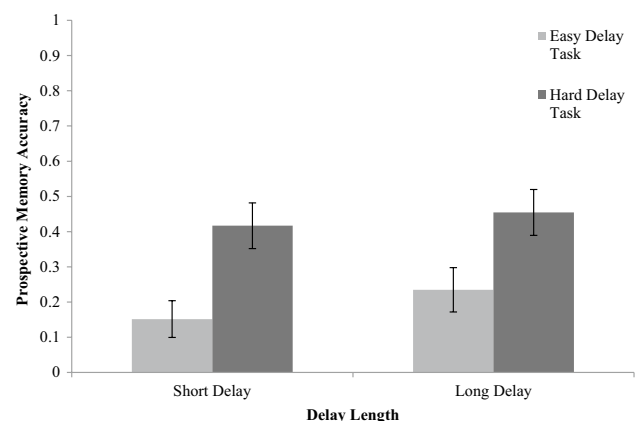


Fig. 1 Prospective memory performance by delay length and delay task difficulty

MSE=0.13, $p=.71$, $\eta_p^2 = 0.001$. Importantly, the main effect of delay task difficulty on PM performance remained significant after controlling for participant's self-reported interest in the delay task as well as their level of tiredness, $F(1, 126)=14.46$, MSE=0.13, $p<.001$, $\eta_p^2 = 0.103$, suggesting that the delay task difficulty effect was not due to differential interest or engagement in the easy and difficult delay tasks.

A 2 (Delay length: short vs. long) by 2 (Delay Task Difficulty: hard vs. easy) ANOVA on PM reaction times for correct responses only revealed no main effect of delay length, $F(1, 60)=0.13$, MSE=266277.10, $p=.72$, $\eta_p^2 = 0.002$, delay task difficulty, $F(1, 60)=0.17$, MSE=266277.10, $p=.68$, $\eta_p^2 = 0.003$, or an interaction between the two, $F(1, 60)=2.66$, MSE=266277.10, $p=.11$, $\eta_p^2 = 0.04$.

Prospective memory and thinking about one's intentions during the delay period

Individuals who reported thinking about the PM task during the delay interval had better PM performance, $r(132)=0.494$, $p<.001$. For individuals who reported thinking about the intention at least once, the number of times they reported thinking about the intention during the delay interval was positively related to PM performance, $r(68)=0.392$, $p=.001$. Interestingly, when the easy and difficult delay task conditions were analyzed separately, results revealed that participants were more likely to report thinking about the PM intention in the difficult delay task ($M=0.65$, $SD=0.48$) than participants in the easy delay task conditions ($M=0.33$, $SD=0.48$; $t(130)=3.83$, $p<.001$, Cohen's $d=0.67$). Further, participants thought about the PM intention *more often* during the difficult delay task ($M=3.26$, $SD=2.24$) than participants in the easy delay task conditions ($M=1.64$, $SD=1.55$; $t(66)=3.19$, $p=.002$, Cohen's $d=0.84$). Participants who reported having the intention always on their mind also did better on the PM task, $r(131)=0.426$, $p<.001$, but there was no difference in participants reporting this in the easy or difficult delay task conditions.

Given that thinking about one's intentions during the delay period was positively related to later PM accuracy and was reported more often in the difficult delay task conditions, we further investigated the role of thinking about the PM intention during the delay task to see if it could account for the main effect of delay task difficulty on PM performance. If thinking about one's intentions was driving the main effect of delay task difficulty (that is, if the difficult delay task, perhaps, provided more opportunities to think about one's intentions compared to the easy delay task), one would expect that how much an individual thought about the PM intention might account for the main effect of delay task difficulty. A 2 (Delay Length:

short vs. long) by 2 (Delay Task Difficulty: hard vs. easy) ANCOVA on PM accuracy with the number of times the individual thought about the PM intention included as a covariate revealed that: (1) how many times an individual thought about the PM intention was a significant covariate of PM performance, $F(1, 63)=8.52$, MSE=0.10, $p=.005$, $\eta_p^2 = 0.119$ and (2) including this covariate in the analysis resulted in the main effect of delay task difficulty becoming non-significant, $F(1, 63)=1.49$, MSE=0.10, $p=.226$, $\eta_p^2 = 0.023$ (an effect size reduction of 78.7%). These results suggest that the extent to which an individual thought about their intention during the delay accounted for the effect of delay task difficulty on later PM performance. Interestingly, whether an individual always had the prospective intention present in their mind during the delay was also a significant covariate of PM performance, $F(1, 126)=28.04$, MSE=0.10, $p<.001$, $\eta_p^2 = 0.182$, but the effect of delay task difficulty remained significant when it was included in the analysis, $F(1, 126)=14.09$, MSE=0.10, $p<.001$, $\eta_p^2 = 0.101$, suggesting that whether an intention was always present in someone's mind did not account for the effect of delay task difficulty on PM performance.

Individuals in the difficult delay task conditions rated the delay task as significantly more interesting ($M=5.68$, $SD=2.03$) than those in the easy delay task condition ($M=4.46$, $SD=2.37$; $t(130)=3.20$, $p=.002$, Cohen's $d=0.55$). Importantly, however, self-reported tiredness, interest in the delay task, and interest in the lexical decision task were unrelated to PM performance, $r_s(132)<0.06$, $p_s>0.286$. Thus, we can rule out some alternative explanations for our main effect of delay task difficulty including differences in fatigue level and motivation/interest in the delay and ongoing tasks.

Discussion

This study examined the impact of the length of delay and delay task difficulty on young adults' PM performance. First, it was confirmed that participants were less accurate as well as slower on the Raven's Matrices (difficult delay task) compared to the item categorization task (easy delay task), showing that our manipulation of delay task difficulty was successful. Our findings showed that a more difficult delay task resulted in better PM than an easy delay task and that this effect was driven by the number of instances that individuals thought about the PM intention during the delay period (and not by always having the intention present in their mind). Our difficult delay task seemed to provide participants with more opportunities to think about their PM intention because of its slower pace, whereas the easy task demanded constant vigilance and immediate responding. There was no effect of delay length and no interaction

between delay length and delay task difficulty on PM performance. Further, tiredness, interest in the delay or ongoing tasks, and having the intention always present in one's mind did not have an impact on the main effect of delay task difficulty, whereas how often a participant thought about their intention during the delay period accounted for the main effect of delay task difficulty.

Curiously, our results surrounding delay length did not replicate Hicks, Marsh, and Russell (2000), as we found no significant effect of delay length on PM performance consistent with some of the literature (e.g., Einstein et al. 1992; Guynn et al. 1998; Stone et al. 2001). One possibility for the difference in our findings is the content of the delay interval. In Experiment 3, Hicks, Marsh, and Russell's (2000) participants did a vocabulary distractor activity where they had to solve synonym and anonym problems, whereas in our delay task, participants either had to solve Raven's Matrices or complete a simple item categorization. It is possible that something about the verbal nature of the delay task in Hicks and colleagues helped to boost PM performance after a longer delay, perhaps, because participants were more primed to think about words and their meanings, whereas our tasks did not include this verbal component. Given the significant impact of delay task difficulty, it is not surprising that using different delay tasks has resulted in different findings between studies. This emphasizes the broader implication of our current study that the content of the task that fills the delay interval has a powerful impact on later PM and thus should be considered carefully when designing PM studies.

Our results support the idea that the activity or task that fills the delay interval is not trivial, but rather may have a meaningful impact on later PM. This is in line with the idea that individuals may use the delay interval to think about their intentions (Hicks et al. 2000) also suggested by our current data that specifically show a positive relation between thinking about the PM intention during the delay and later PM performance. Further, our results challenge the assumption that a cognitively demanding task necessarily provides fewer opportunities to think about one's future intentions. Our results contrast with the previous findings that suggest that delay task difficulty has no impact on later PM (Cook et al. 2014; Shelton et al. 2011, 2013), but differences between findings might be accounted for by the use of different delay tasks and the availability of opportunities to think about the prospective intention. Previous work with preschool-aged children has shown that a difficult delay task reduced later PM performance (Mahy and Moses 2015); however, the delay task in that particular study required constant attention and vigilance, as it was a working memory task. In contrast, our current results suggest that delay tasks that allow more time for pauses and reflection, such as the Raven's Matrices, may actually

promote mind wandering and opportunities to refresh one's intentions compared to relatively simple tasks, such as item categorization tasks, that require fast responses with little time to refresh intentions within or between trials. In line with this rationale, participants took over 7 s on average to solve a difficult delay task trial (Raven's Matrices), whereas they took less than 1 s to solve an easy delay task trial (item categorization task). It is likely that in the current study, individuals had a little opportunity to think about their intentions during the easy item categorization task as they performed on average around 84 trials per minute. Finally, our results ruled out alternative explanations for the delay task difficulty effect, so that differences in later PM performance cannot be attributed to differences in levels of tiredness or interest in the delay task as once these factors were controlled for the effect of delay task difficulty persisted. Notably, although individuals reported the Raven's Matrices was more interesting than the item categorization task, they seemed to be more likely to shift their attention away from the task to think about the PM task. This is likely due to the slower pace of the Raven's Matrices that allowed for mind wandering and refreshing intentions, perhaps, especially in the challenging matrices that took longer to solve.

An alternative explanation for superior PM performance after a difficult delay task was that the easy delay task, which required more verbal processing to categorize items into living and non-living things, might have interfered with the PM task due to its verbal nature (and thus resulted in lower PM performance). In contrast, the difficult delay task relied on visuospatial reasoning to complete Raven's matrices that might not have interfered with later verbal processes involved in detecting the PM cue words. Future work should examine whether domain-specific vs. domain-general cognitive load discourages self-reminding equally and what impact this might have on PM performance.

Again, although a longer delay has been suggested to provide more opportunities to think about one's intentions in past research (Hicks et al. 2000), we did not find any effect of delay length or interaction with delay task difficulty. Any delay length effect seems to be overshadowed in the current study by the demands of the delay task, where a task that keeps a participant constantly busy seems to limit opportunities to think about one's intentions and had a negative effect on PM performance even over a longer delay period of 15 min. Thus, difficulty of the delay task seems to have a more powerful impact on PM performance than delay length alone and does not interact with delay length. Interestingly, it seems that studies where everyday life activities fill the delay (i.e., naturalistic studies) tend to find negative effects of delay length on PM. Future work could examine delay length (as well as delay task difficulty) in a meta-analysis to clearly establish which effects hold up across studies and which do not.

Individuals who reported thinking about the intention more in the delay interval were also more likely to carry out their intentions at the appropriate time similar to Kvavilashvili and Fisher's (2007) findings. In a naturalistic study, Sellen, Louie, Harris, and Wilkins (1997) showed that there was an increase in thinking about the PM intention in everyday contexts during natural transition periods between tasks. It is possible that the Raven's Matrices provided several natural transitions either between problems or when an individual took a mental break from a problem and thus may have promoted thinking about their PM intention. Past studies that have included breaks during the delay period between tasks have shown that these breaks promote later PM performance (Hicks et al. 2000) or benefit later PM when individuals are explicitly instructed to think about the PM intention during these breaks (Finstad et al. 2006). In contrast, our results suggest that even if individuals are not explicitly told to think about the PM intention they do so during the delay period, especially when there is a slower paced albeit difficult task that allows for reflection on the PM intention. Future studies should consider longer delays than 15 min that might better approximate everyday life (e.g., delays of 1 h and more) and should also attempt to examine individual differences in self-reflection or thinking about intentions as it was shown to relate to PM performance in the current study. Further, more work is needed to more thoroughly examine the different effects that manipulating features of the delay task (beyond task difficulty) have on later PM performance, as often in the literature little attention is paid to the delay period activity and participants fill out questionnaires or complete other tasks without considering that these tasks might impact later PM performance.

A further contribution of the current study is the finding that individuals seem to have substantial insight into internal mental processes (such as thinking about a future intention) and self-reports of how much an individual engages in thinking about the intention during the delay was positively related to later PM performance. Thus, self-reports are a useful source of information and reveal that adults who reflect on their intentions during the delay interval are more likely to carry them out at a later point in time. Future research would do well to continue to ask participants about their mental processes that occurred during the delay or ongoing tasks to capture more information about the mechanisms supporting successful PM performance.

A limitation of the current study was that our manipulation of task difficulty used two different tasks (Raven's Matrices and living/non-living judgment task) rather than a single task with two levels of difficulty. Future work should extend our results to examine whether delay task difficulty manipulated within the same task has the same impact on PM performance and refreshing one's

intentions. Further, although our post-experimental assessments revealed interesting relations between how often individuals thought about their intention and later PM performance, it is possible that insights into mental activity could have been influenced by PM accuracy (i.e., individuals who did well on the PM task might have reported thinking about the intention more during the delay period). Thus, future work should examine how often an individual is thinking about their intention in real time rather than relying on retrospective reports after the PM task.

Our current results have methodological implications for future studies. Specifically, attention should be paid to what tasks are being completed during the delay interval rather than conceptualizing the delay interval as a chance to measure multiple secondary abilities. Further, while the delay interval was originally considered an interval in which to promote forgetting of the PM intention, it is clear that much more is occurring during this phase. In fact, the current study shows that the demands of the delay task have an impact on later PM performance and that thinking about one's PM intentions during this period is positively correlate with later PM performance. While much attention has focused on the effect of the length of delay on later PM performance (e.g., Brandimonte and Passolunghi 1994; Hicks et al. 2000; Nigro and Cicogna 2000), our results suggest that the content of the delay interval is just as important, if not more important, than delay length. Future work should continue to examine the relations between task difficulty, delay interval length, and PM performance, as this is an area of research that is potentially complex and merits further exploration.

Acknowledgements Preparation of the manuscript was partially funded by a Natural Sciences and Engineering Council of Canada Discovery Grant (RGPIN-2015-03774) and Swiss Government Scholarship to CEVM. MK acknowledges funding from the Swiss National Science Foundation (SNSF). The authors wish to thank Chirine Ajram, Joelle Barthassat, Katelyn Brausewetter, Riley Brennan, Malik Djela, Alison O'Connor, and Delphine Paumier for their assistance with data collection.

Compliance with ethical standards

Funding This study was funded by the Natural Sciences and Engineering Council of Canada (Grant Number: RGPIN-2015-03774 to Caitlin Mahy).

Conflict of interest Caitlin Mahy, Katharina Schnitzspahn, Alexandra Hering, Jacqueline Pagobo, and Matthias Kliegel declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

Appendix

Participant Questionnaire

Forced choice responses are indicated in bold.

1. Please describe what you had to do in the lexical decision task. *If participant did not describe the lexical decision task rules and the rule of pressing the '9' key when they saw an animal word, the experimenter followed-up with probes: (1) What were you supposed to do in the lexical decision task? (2) Was there something else you had to do in the lexical decision task? (3) What were you supposed to do when you saw an animal word? (Administered verbally by the experimenter).*
2. Did you think about pressing the button for an animal word during the problem-solving/ item categorization task? YES / NO.
 - If so, how many times?
3. Was the intention to press a button when you saw an animal word always present in your mind? YES / NO
4. On a scale of 1 (not at all) to 10 (extremely), how tired are you right now?
5. On a scale of 1 (not at all) to 10 (extremely), how interesting was the problem-solving/ item categorization task?
6. On a scale of 1 (not at all) to 10 (extremely), how interesting did you find the lexical decision task?
7. How old are you?
8. What is your birth date?
9. What is your biological sex? MALE / FEMALE
10. What is your academic major?
11. What is your ethnic background? Please check all that apply.
 - White
 - Black or African American
 - Hispanic, Latino, or Spanish
 - Asian
 - Asian Indian
 - Hawaiian Native
 - Pacific Islander
 - Middle Eastern

- Alaskan Native or American Indian
- Other group (please specify) : _____.

References

- Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*, 445–459.
- Brandimonte, M. A., & Passolunghi, M. C. (1994). The effect of cue-familiarity, cue-distinctiveness, and retention interval on prospective remembering. *The Quarterly Journal of Experimental Psychology*, *47*, 565–587.
- Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, *10*, 12–21.
- Cook, C., Ball, B. H., & Brewer, G. A. (2014). No effects of executive control depletion on prospective memory retrieval processes. *Consciousness and Cognition*, *27*, 121–128.
- Ebbinghaus, H. (1885). *Über das gedächtnis: Untersuchungen zur experimentellen psychologie*. Duncker & Humblot.
- Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Dover. (Original work published 1885).
- Einstein, G. O., Holland, L. J., McDaniel, M. A., & Guynn, M. J. (1992). Age-related deficits in prospective memory: The influence of task complexity. *Psychology and Aging*, *7*, 471–478.
- Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 717–726.
- Finstad, K., Bink, M., McDaniel, M., & Einstein, G. O. (2006). Breaks and task switches in prospective memory. *Applied Cognitive Psychology*, *20*, 705–712.
- Guynn, M. J., McDaniel, M. A., & Einstein, G. O. (1998). Prospective memory: When reminders fail. *Memory & Cognition*, *26*, 287–298.
- Hicks, J. L., Marsh, R. L., & Russell, E. J. (2000). The properties of retention intervals and their affect on retaining prospective memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1160–1169.
- Johansson, O., Andersson, J., & Ronnberg, J. (2000). Do elderly couples have a better prospective memory than other elderly people when they collaborate? *Applied Cognitive Psychology*, *14*, 121–133.
- Kvavilashvili, L., & Fisher, L. (2007). Is time-based prospective remembering mediated by self-initiated rehearsals? Role of incidental cues, ongoing activity, age, and motivation. *Journal of Experimental Psychology: General*, *136*, 112–132.
- Mahy, C. E. V., & Moses, L. J. (2015). The effect of retention interval task difficulty on children's prospective memory: Testing the intention monitoring hypothesis. *Journal of Cognition and Development*, *16*, 742–758.
- Mahy, C. E. V., Moses, L. J., & Kliegel, M. (2014). The development of prospective memory in children: An executive framework. *Developmental Review*, *34*, 305–326.
- Mahy, C. E.V., & Moses, L. J. (2011). Executive functioning and prospective memory in young children. *Cognitive Development*, *26*, 269–281.
- Martin, B. A., Brown, N. L., & Hicks, J. L. (2011). Ongoing task delays affect prospective memory more powerfully than filler task delays. *Canadian Journal of Experimental Psychology*, *65*, 48–56.

- McBride, D. M., Coane, J. H., Drwal, J., & LaRose, S. A. M. (2013). Differential effects of delay on time-based prospective memory in younger and older adults. *Aging, Neuropsychology, and Cognition*, *20*, 700–721.
- Nigro, G., & Cicogna, P. C. (2000). Does delay affect prospective memory performance? *European Psychologist*, *5*, 228–233.
- Nigro, G., Senese, V. P., Natullo, O., & Sergi, I. (2002). Preliminary remarks on type of task and delay in children's prospective memory. *Perceptual and Motor Skills*, *95*, 515–519.
- Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13.
- Raven, J. C. (1941). Standardization of progressive matrices. *British Journal of Medical Psychology*, *19*, 137–150.
- Scullin, M. K., & McDaniel, M. A. (2010). Remembering to execute a goal sleep on it! *Psychological Science*, *21*, 1028–1035.
- Sellen, A. J., Louie, G., Harris, J. E., & Wilkins, A. J. (1997). What brings intentions to mind? An in situ study of prospective memory. *Memory (Hove, England)*, *5*, 483–507.
- Shelton, J. T., Cahill, M. J., Mullet, H. G., Scullin, M. K., Einstein, G. O., & McDaniel, M. A. (2013). Resource depletion does not influence prospective memory in college students. *Consciousness and Cognition*, *22*, 1223–1230.
- Shelton, J. T., McDaniel, M. A., Scullin, M. K., Cahill, M. J., Singer, J. S., & Einstein, G. O. (2011). Cognitive exertion and subsequent intention execution in older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *66*, 43–50.
- Somerville, S. C., Wellman, H. M., & Cultice, J. C. (1983). Young children's deliberate reminding. *The Journal of Genetic Psychology*, *143*, 87–96.
- Souza, A. S., Rerko, L., & Oberauer, K. (2014). Unloading and reloading working memory: Attending to one item frees capacity. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1237–1256.
- Stone, M., Dismukes, K., & Remington, R. (2001). Prospective memory in dynamic environments: Effects of load, delay, and phonological rehearsal. *Memory (Hove, England)*, *9*, 165–176.
- Szarras, K., & Niedźwieńska, A. (2011). The role of rehearsals in self-generated prospective memory tasks. *International Journal of Psychology*, *46*, 346–353.
- Tierney, S.M., Bucks, R.S., Weinborn, M., Hodgson, E., & Woods, S.P. (2016). Retrieval cue and delay interval influence the relationship between prospective memory and activities of daily living in older adults. *Journal of Clinical and Experimental Neuropsychology*, *38*, 572–584.