Explaining Complex Age-Effects on Prospective Memory

Simon J. Haines

A thesis by publication submitted in total fulfilment of the criteria for the degree of

Doctor of Philosophy

School of Psychology

Faculty of Health Sciences

Australian Catholic University

March 2019

Declaration

This thesis contains no material that has been extracted in whole or in part from a thesis that I have submitted towards the award of any other degree or diploma in any other tertiary institution.

No other person's work has been used without due acknowledgment in the main text of the thesis.

All research procedures reported in the thesis received the approval of the relevant Ethics/Safety Committees (where required).

[redacted]

Simon Haines

PhD Candidate

Date 22/3/2019

Acknowledgements

There are many people who have supported me during my PhD candidature and in a multitude of ways. Peter Rendell, my primary supervisor, has provided invaluable guidance and been very generous in introducing and connecting me with many brilliant researchers in the PM field. Two of these wonderful researchers and people I have had the privilege of having as co-authors and guides are Jill Shelton and Julie Henry. My secondary supervisor, Lucy Busija has offered more guidance and wisdom on statistical analyses and critical thinking than I could have ever dreamed possible. I owe a great deal to Lucy and thank her for her patience. Skye McLennan has also been a wonderful support, particularly during the grueling data collection phase and has been a reliable and acute commentator on my written work. I would also like to warmly thank Colleen Doyle, Gemma Tatangelo, Gill Terrett, Alexandra Hering, Peter Koval, Izelle Labuschagne, Nate Rose, and Peter Goodin, who have all been excellent academic and research guides. From my cohort I would like to thank Amy Malcom, Julian Rossi, Caitlin Grace, Serene Chua, Grace Hayes, Rachel Pelly, Adam Clemente, and Phoebe Imms. I would especially like to acknowledge Nick Koleits, for heroic support during data collection and for IT trouble shooting; and Jenny Matthews and Bernardo Jarrin, for cheerfully helping me with all things admin related. Kathryn Duncan, senior librarian at ACU, also gave me vital support in the final stages of the thesis. Necessary for the very possibility of my current research, I would also like to acknowledge and express my gratitude to all the participants who took part in the Memtrain study. Finally, but by no means least, I would like to thank my family, all of them, but most importantly, for my sanity—then, now and in the future—I want to thank and dedicate this thesis to my partner Zohre and my daughter Zerya.

Table of Contents

Contents

List of Tables

List of Figures

A note on spelling

For accepted and submitted for publication manuscripts (Chapters 2 and 4) United States English has been used as required by the relevant journals or outlet editors. In other sections of thesis Australian English has been used, for example “ageing” rather than “aging”.

Abstract

The general aim of this thesis was to investigate complex age-related differences in prospective memory (PM). An extensive review of PM and cognitive ageing (an accepted peer reviewed article for *Oxford Research Encyclopedias*) was the starting point for identifying key areas of further research in the age-PM field. Specifically, two areas of further research were identified and then investigated empirically in the present research project. First, the need to use conceptually parallel PM task types across settings to illuminate mechanisms of the age-PM paradox was identified. The paradox refers to the general finding that young outperform older adults in the laboratory, but vice versa in naturalistic-settings; and young-old outperform old-old in laboratory but show equivalent levels of performance in naturalistic-setting studies. A second area identified as requiring further empirical investigation was the role of executive functions in mediating age-effects across a wide range of PM tasks. The current research project made a significant contribution to the age-PM literature by undertaking a series of empirical investigations into both of these areas.

The first key empirical investigation is related to illuminating the mechanisms of the age-PM paradox. A paper reporting two empirical experiments is presented (which have been *submitted to the Journal of Experimental Psychology: General*). In Experiment 1, young (19–30 years; *n* = 40) and older adults (65–86 years; *n* = 53) were tested on conceptually parallel PM tasks in the laboratory (using the Virtual Week paradigm that simulates naturalistic PM tasks and routine daily activities) and in a naturalistic-setting using a recently developed novel smartphone paradigm, the MEMO. PM tasks were conceptually parallel in terms of the type of cue and inherent level of environmental support afforded by three PM task types: event-based, time-of-day, and time-interval. The latter two time-based tasks were hypothesised to largely account for

the age-PM paradox, in particular by not being sufficiently distinguished and investigated separately in previous studies of the age-PM association in performance across settings. In Experiment 1, participants completed two simulated days of Virtual Week with the three PM tasks types embedded, and two *separate* blocks of three days for the event-based PM (a photo task when particular events were encountered) and two time-based PM tasks (scheduled and "pop up" quizzes, e.g., come back in 10 minutes to open app on phone to complete a quiz). It was found that young adults outperformed older adults in the laboratory, however, in the naturalistic-setting older adults only outperformed younger adults on time-of-day tasks (i.e., appointment like tasks; in this case completing a scheduled quiz), while similar levels of performance were revealed on event-based tasks (relatively high performance) and time-interval tasks (relatively low performance). In Experiment 2 young-old (60–74 years; n = 64) and old-old (75–87 years; n = 40) adults were compared on the same two PM task types. However, this time the naturalistic-setting paradigm, MEMO, was made more challenging, and conceptually closer to Virtual Week, by combining all PM task types over the same three day period. Participants were also permitted to use external aids. The results showed that young-old outperformed old-old in the laboratory, with both age groups showing better performance on event-based than the time-interval tasks (involving monitoring a stop clock in Virtual Week). Together, these experiments show that the age-PM paradox is only apparent when comparing different types of time-based tasks across settings, and that older adults are vulnerable to forgetting delayed intentions over short intervals with relatively few time cues in a naturalistic-setting. However, when permitted to use external aids older adults can compensate for this cognitive vulnerability, and show similar high levels of performance to their own performance on event-based and time-of-day tasks.

The second key empirical investigation is related to the generally hypothesised mediating role of facets of executive functioning on age-effects on PM, particularly for putatively high demand tasks. In a very novel study in the age-PM field, for the first time four different PM measures were combined in a single study to investigate individual differences in facets of executive functioning mediating age-effects on PM. The study used a large sample of older adults (*n* = 104; range 60–87 years) who performed the laboratory paradigm Virtual Week, two clinical measures, the Cambridge Prospective Memory Test (CAMPROMT), and the Memory for Intentions Test (MIST), and the MEMO. The results showed that there were age-effects for the PM measures that presumably had the highest cognitive demands in so far as they do not permit external aids, that is Virtual Week and the MIST (external aids are allowed in the CAMPROMT and in the present study's use of the MEMO paradigm). Contrary to previous studies using abstract PM paradigms in the laboratory, older adults' performance on these relatively naturalistic (even familiar) PM task types showed some relationship to retrospective memory processes, processing speed, and age, but virtually no relationship to separable facets of executive functioning. The models tested included both parallel mediation models, in which both executive and non-executive cognitive processes were measured, and a serial mediation model with processing speed hypothesised to impact executive processes which in turn influence PM task performance. The lack of evidence for both these models suggest that PM is a relatively independent functional construct that is affected by age in some circumstances, regardless of individual differences in executive functioning, and is also sometimes independent of both age and executive functioning processes. The present thesis thus indicates potential boundary conditions for both the manifestation of the age-PM paradox and the mediating role of executive functions on age-related changes in PM.

Chapter 1: General Introduction

The view that prospective memory should be especially difficult for the elderly rests on the questionable assumption that all prospective memory tasks are alike in that they are high in self-initiated retrieval.

– Einstein and McDaniel (1990, p. 724)

At the between-tasks level, a theory of cognitive aging should provide an account of the types of tasks that should demonstrate evidence of age-related decline and those which are age insensitive. At the within-task level, a theory of cognitive aging must be able to distinguish those task conditions that produce age-related deficits from those that do not produce age-related deficits.

– West (1996, p. 283)

1.1 Overview

The study of prospective memory (PM), the complex memory ability of recalling and executing future intentions at the right time or when indirectly cued by an event (Einstein & McDaniel, 1990; McDaniel & Einstein, 2007), has had a long association with concerns in the cognitive ageing field (Craik, 1986; Einstein & McDaniel, 1990; Moscovitch, 1982). The existence and extent of age-related declines in PM has also been contested from both apparently conflicting empirical findings (Cherry & Lecompte, 1999; d'Ydewalle, Luwel, & Brunfaut, 1999; Einstein & McDaniel, 1990; Rendell & Thomson, 1999) and the theories that have interpreted this research (McDaniel & Einstein, 2000; Smith & Bayen, 2004) and shaped subsequent research and theoretical developments (Shelton & Scullin, 2017). The first chapter of this thesis is designed to orientate the reader in this important field of research (Brandimonte, Einstein, & McDaniel, 1996; Ellis & Kvavilashvili, 2000; Harris & Wilkins, 1982; Kliegel, McDaniel, &

Einstein, 2008). Including introducing the key PM task types: *event-based* tasks, where the cue to execute a delayed intention is the occurrence of a particular event (e.g., passing on a message when seeing a colleague); and *time-based* tasks, where the cue to execute a delayed intention is either a *time-of-day* (e.g., dentist appointment at 3pm) or the elapsing of a specific *time-interval* (e.g., move the car in two hours to avoid a fine).

The second chapter is a comprehensive literature review of ageing and PM (Haines et al., 2019). It introduces the key past and current research findings, and the main theories and their developments up to the present. As a standalone encyclopedia entry for the *Oxford Encyclopedia of Psychology and Aging*, it also provides a comprehensive starting point for anyone interested in the age-PM research field and what further work is required. Some of the further work required, flagged in this literature review, is empirically addressed by the present thesis. The third chapter is a detailed description of the general methodology used in this thesis. The fourth chapter, submitted for publication, comprises two studies investigating the age-PM paradox—the paradoxical finding that older adults have poor PM performance relative to younger adults in the laboratory, but vice versa in naturalistic-settings (Henry, MacLeod, Phillips, & Crawford, 2004; Rendell & Craik, 2000; Rendell & Thomson, 1999); *and* the relatively neglected second aspect of a parallel difference between young-old and old-old adults in the laboratory (young-old superior to old-old) but no difference in naturalistic-settings (Kliegel, Rendell, & Altgassen, 2008; Rendell & Craik, 2000). This research was carried out with a relatively new, naturalistic-setting, PM paradigm (MEMO; Randall, 2016)[1]; with PM task types designed to match those typically used in an established laboratory paradigm (i.e., Virtual Week; Rendell & Craik, 2000).

[1] To avoid confusion it should be noted that the moniker "MEMO" is not an acronym. To date, the MEMO has only appeared in an unpublished PhD thesis manuscript (Randall, 2016).

An empirical study addressing the question of the putative role of facets of executive functioning ability in understanding age-effects on PM is presented in the fifth chapter. The findings of this study indicate that executive functioning ability does not necessarily mediate age-effects on PM performance for all PM task types and paradigms, thus offering some novel insights and challenges for the current theoretical approaches that attempt to explain age-related decline in PM. The theoretical implications of this work, along with consideration of methodological issues of PM assessment, and how this thesis's empirical findings compare and fit within the ageing-PM research field are discussed in the sixth, final, chapter. This last chapter also highlights the unique and substantial contribution made by the research undertaken and presented in this thesis, as well as indicating directions for future research.

1.2 The Current Research Project

The general aim of the current research project is to shed light on and help to resolve some of the pressing concerns in the field of prospective memory (PM) and cognitive ageing. These concerns include understanding how PM task characteristics interact with cognitive changes associated with ageing, and how individual differences in cognitive processes within older adults may mediate patterns of PM performance found in both laboratory and naturalistic-settings. To facilitate greater clarity on the current research project's goals, the specific aims of this research are set out in a schematic format.

1.1.1 Specific aim one

A key concern of the current research project is to breach the gap and explain the discrepant (even paradoxical) findings of age differences in PM performance across laboratory and naturalistic-settings. The age-PM paradox is the surprisingly robust pattern in the age-PM literature of young adults generally outperforming older adults in the laboratory, while in naturalistic-settings the pattern is reversed such that older adults consistently outperform younger adults. This pattern was found even when a complex time-logging regimen (designed to simulate medication taking) was given to young and older adults in a naturalistic-setting (Rendell & Thomson, 1999). The same study by Rendell and Thomson (1999) also demonstrated a second aspect of the PM paradox, namely that young-old adults consistently outperform old-old adults on PM tasks in the laboratory, while both older adult groups perform at comparable levels on PM tasks in naturalistic-settings. In addressing this aim, the present research builds on the early work on developing parallel PM tasks across laboratory and naturalistic-settings carried out by Rendell and Craik (2000) and, more recently, Randall (2016). Specifically, the current research project used the Virtual Week laboratory paradigm (Rendell & Craik, 2000; Rendell & Henry, 2009) that simulates a daily life narrative, with different types of everyday PM tasks embedded, using a computerised board game format. Conceptually parallel to the Virtual Week in terms of PM task types, the MEMO paradigm is carried out using a mobile phone camera and customised application. The balance of ecological validity and high level of control, afforded by both Virtual Week and MEMO, enables a comprehensive investigation of the mechanisms contributing to the age-PM paradox and a refinement of the theoretical models of interactions between age and PM task type characteristics.

1.1.2 Specific aim two

A second specific aim of this research project is to focus on age-effects and PM within a healthy older adult cohort. Given that most cognitive decline associated with ageing tends to occur in the later decades of older adulthood life (Huppert, Johnson, & Nickson, 2000; Kliegel & Jager, 2006; Park & Reuter-Lorenz, 2009) it is important to consider not just young adults compared to older adults on PM tasks, but also the differences across laboratory and naturalistic-settings between young-old (typically 60-74 years of age) and old-old (typically 75+ years of age) adults.[2] Early research by Einstein and McDaniel (1990) suggested that PM may be a cognitive function that is relatively spared the detrimental effects of ageing found in other cognitive processes, such as processing speed (Verhaeghen, 2014) and retrospective memory (Craik, 1986). However, the boundary conditions for spared PM in later life may depend on the nature of the PM task and the cognitive processes underpinning PM performance on that task (McDaniel & Einstein, 2000).

1.1.3 Specific aim three

A third specific aim of this research project is to investigate how PM task type characteristics, in particular level of environmental support, make differing demands on the cognitive processes theorised to mediate age-effects in PM performance. A key development, discussed in the literature review (chapter 2) is that PM tasks are not all equal in self-initiated processing demands (McDaniel & Einstein, 2000). Instead there is a broad distinction between PM tasks. Some are relatively high in environmental support, such as when a salient event acts as the cue to execute the PM task, whereas others are relatively low in environmental support, such as when a time-interval elapsing (e.g., the end of a 10 minute break at work) with no salient cue determines when the PM task is to be executed. In the latter, time-based PM task, self-initiated processes such as monitoring and switching attention (e.g., intermittently checking the time) are required; whereas with the former, event-based task, relatively spontaneous processes can potentially be relied on to retrieve the intention (McDaniel & Einstein, 2000).

[2] The group division between young-old and old-old is relatively arbitrary given that the manifest cognitive ageing process is not uniform across all older individuals (Park & Reuter-Lorenz, 2009), however, here we follow previous work in the literature (Huppert et al., 2000) in making this general distinction for research purposes. In chapter 5 this distinction is not made and an individual differences approach within older adults is adopted when investigating the mediating role of facets of executive functioning in PM performance.

1.1.4 Specific aim four

The fourth specific aim of this research project is to highlight, formalise, investigate, and promote research on age-effects in terms of a relatively neglected distinction within time-based PM tasks: namely, time-of-day versus time-interval tasks. The broad distinction between event and time-based PM task types permits of finer distinctions within each task type. Event-based PM tasks are often divided into focal and non-focal tasks, where the former involves a cue for executing the PM task which is related to the cognitive processing involved in performing the ongoing or “distractor” task, while the latter has an unrelated cue. Focal tasks tend to rely primarily on spontaneous retrieval processes (similar to recognition based memory tasks) which are relatively immune from age related decline, while non-focal tasks tend to rely on controlled attentional processes which typically decline with age. The focal versus non-focal distinction is almost exclusively applied to event-based cues. However, there are hints in the age-PM literature that a similar distinction can be made in regards to time-based tasks (McDaniel & Einstein, 2007; Rendell & Craik, 2000), and that by making an analogous distinction the discrepant findings of older adults’ performance across laboratory and naturalistic-settings might be explained (at least partially).

Some naturalistic-setting time-based tasks (i.e., time-of-day tasks) have environmental cues which are similar in terms of environmental support to the relational cues of focal event-based tasks. For example, attending a medical appointment at 4pm in the afternoon may include the relational cues of position of the sun, quantity of traffic, or favourite radio programs, all of which may be more or less related to an ongoing task (e.g., walking or driving through the city) and lead to spontaneous retrieval of the intention to go to the appointment. In contrast, other time-based tasks (i.e., time-interval tasks) may be analogous to non-focal event based tasks in so far as the task appropriate processing of the ongoing task (e.g., reading and responding to emails) is unrelated to, or lacks environmental, time-based, cues to retrieve and execute, the PM task (e.g., returning a phone call after ten minutes has elapsed). The key point that this thesis aims to make on this score is that it is erroneous to assume that time-based tasks are homogenous in their cognitive demands of limited cognitive processing resources, just as in the case of event-based PM tasks.

In the following chapters it is argued theoretically, and investigated empirically, that there are two types of time based PM tasks: time-of-day tasks and time-interval tasks, analogous to event-based focal and non-focal PM tasks in terms of cognitive demand and potential age-effects. Time-of-day time-based PM tasks are appointment-like tasks, which are best illustrated in a naturalistic-setting, where there is a rich array of time-cues that may interact with both spontaneous retrieval of the intention, and with monitoring when the right constellation of contextual cues are present (Rabbitt, 1996). A challenge for laboratory paradigms is to replicate (or simulate) this rich form of environmental support in a controlled setting. One paradigm that aims to do this is Virtual Week (Rendell & Craik, 2000; Rendell & Henry, 2009), a laboratory paradigm that stimulates the kind of daily life narratives and routines that likely characterise many older adults' lives and provides scaffolding for compensating for cognitive decline and executing PM tasks on time. The other type of time-based task, the time-interval task, has dominated laboratory studies, but been conspicuously absent from naturalistic-setting studies, partly, it is argued, because of a lack of theoretical appreciation of this time-based task distinction. In time-interval tasks, a participant needs to perform a task after a certain time-period has elapsed (e.g., calling someone back in 15 minutes; Rose et al., 2015). The time-interval task can appear deceptively easy. However, there is strong evidence from laboratory studies that this is a very difficult task, particularly for older adults. In both of the empirical investigations in the current research project, this distinction between time-of-day and time-interval tasks is made explicit, and age-effects on each are measured.

Two research paradigms that have conceptually parallel PM task types are used in this research project. The first is the well-established laboratory paradigm Virtual Week (Rendell & Craik, 2000; Rendell & Henry, 2009; Rose et al., 2015; Terrett et al., 2019), which includes event-based tasks, simulated time-of-day tasks, and time-interval tasks. The second paradigm, the MEMO (Randall, 2016), is relatively new but resembles a previous paradigm, Actual Week (Rendell & Craik, 2000), that aimed to conceptually match PM task types in a naturalistic-setting with those used in Virtual Week. The MEMO utilises a smartphone application and camera function allowing the collection of time and date-stamped data to accurately measure performance on each PM task type. Event-based tasks involve the use of the camera function, while the time-based tasks involve participants opening the MEMO application on the phone at either scheduled times of day (which are selected by the participant each morning) or after a short period of time has elapsed (time-interval tasks). The latter time-interval task involves a quasi-random notification via the MEMO application to "come back" after a particular amount of time has elapsed to complete a short quiz. Both time-based tasks involve this short quiz, which consists of three questions capturing contextual information about where, with whom, and what the participant was doing just prior to executing the PM task.

1.1.5 Setting the scene

To set the scene, the current research project includes a comprehensive overview of the research to date on the association of cognitive ageing with PM (Chapter 2). This literature review chapter, which includes a synthesis and critique of theory and empirical research in the field, and highlights the gaps in current research, has been published in the *Oxford Encyclopedia of Psychology* (Haines et al., 2019). It takes both a historical and thematic approach to the theorising and empirical investigation of how ageing relates to PM. It also provides a justification for the empirical research that follows in subsequent chapters. One of these needed areas of research is in continuing to refine parallel PM tasks across laboratory and naturalistic-settings. For example, there is a need to more closely match cognitive load in a naturalistic-setting by *combining* both time and event-based tasks within the same test period (see Experiment 2, Chapter 4). Previously the majority of naturalistic-setting studies (including the previous one which used the MEMO paradigm; Randall, 2016) have only given event-based or time-based tasks separately (though for two naturalistic setting studies that included event and time-of-day PM tasks, and event and time-interval PM tasks, respectively, see Niedźwieńska & Barzykowski, 2012; Sellen, Louie, Harris, & Wilkins, 1997; note that only the first study included older adults). Continuing incremental empirical research on deficits in older adults' PM performance on *diverse* "high demand" PM tasks is also required, and addressed in this research project. In particular, it is investigated whether older adults PM deficits on a range of PM tasks can be traced back to deficits in controlled processing (or facets of executive functioning ability).

The research project was carried out using data collected by the author of this thesis as part of a larger parent study on cognitive training (see Methodology chapter, Chapter 3). It also included some secondary data (Experiment 1, Chapter 4; Randall, 2016). The primary data collected included event-based and time-based PM performance on the Virtual Week laboratory paradigm, the MEMO naturalistic-setting paradigm, and two clinical measures of PM, the Cambridge Prospective Memory Test, and Memory for Intentions Screening Test. There were also measures of facets of executive functioning (shifting, updating, and inhibition; see Chapters

3 and 6), processing speed, and retrospective memory. The details of the measures and procedure for data collection is presented in the Methodology chapter (see Chapter 3).

1.3 Theoretical beginnings

Interest in “forgetfulness” of future intentions—here referred to as PM lapses—has had a long history (Freud, 1901/1952; Lewin, 1926/1999). Freud (1901/1952) incorporated PM into his grand theory of action, analyzing various “specimens” of forgetting, including romantic dates, military duties, and paying bills. According to Freud, the common thread of these forms of forgetting is a murky “counter-will’, or an unconscious lack of motivation to execute the PM task. Although this interpretation may account for some cases of forgetting (e.g., “you forgot flowers for our anniversary because it obviously does not matter much to you anymore”) it does not address or seek to elucidate cognitive vulnerabilities associated with ageing (e.g., decline in attentional processes). Freud (1901/1952) does, however, give an eloquent description of an everyday PM task carried out by someone (himself) who is not plagued with a neurosis:

> If I am going for a walk and take a letter with me which has to be posted, it is certainly not necessary for me, as a normal individual, free from neurosis, to walk all the way with it in my hand and to be continually on the look-out for a letter-box in which to post it; on the contrary I am in the habit of putting it in my pocket, of walking along and letting my thoughts range freely, and I confidently expect that one of the first letter-boxes will catch my attention and cause me to put my hand in my pocket and take out the letter. (p. 152)

This description suggests relatively spontaneous processes of retrieving an intention—which may not be affected by ageing (Einstein & McDaniel, 1990)—when there is a clear cue (e.g., a letter-box) to execute the PM task (e.g. mailing a letter).

Lewin (1926/1999) was perhaps the first psychologist to explicitly acknowledge the difference between prospective and retrospective memory, noting that:

> Two concepts of forgetting must be carefully distinguished. The first pertains to the ability to reproduce knowledge once possessed. The second concept of forgetting pertains to intentions which are not carried out. In everyday life, we call it “forgetfulness.” It is obvious that we usually remember the content of the intention, even though we have forgotten to carry it out. (p. 89)

Lewin observed how a PM cue (e.g., a mail box) or constellation of cues (in the common case when there is no specific cue) tends to be encoded together with a PM task (e.g., mailing a letter). An acute observer, Lewin also noted that the holding of delayed intentions “do not seem to show a time-decrement” (p.90), in contrast to the observation of Ebbinghaus (1885/2013, p. 156) that “Left to itself every mental content gradually loses its capacity for being revived, or at least suffers loss in this regard under the influence of time.” In other words, for Lewin, as for some contemporary researchers of ageing and PM (e.g., Smith, 2003), until the intention is executed (or some appropriate substitution is found; e.g. giving the letter to a friend to mail) there appears to be a cognitive load or tension which is carried without being continually refreshed in consciousness. This hypothesised “inner-tension” of holding a delayed intention is important, given that *cognitive capacity*, how much information can be maintained, appears to decline with age (Kidder, Park, Hertzog, & Morrell, 1997). It suggested that holding one or more intentions may lead to older adults showing “costs” in other tasks competing for their attention, and trade-offs on either PM task performance or performance on other activities—which is exactly what has been found in subsequent experiments (Logie, Maylor, Della Sala, & Smith, 2004; Smith, 2003).

Lewin also observed how interruptions can lead to retrieval failures–the feeling that "there is still something to be done" (p.88)–again, this is particularly relevant to older adults having PM lapses as interruptions constitute a form of cognitive demand, divided attention, that may interact with the general finding that episodic retrospective memory declines with age in general (Craik, 1986). One potential reason that older adults do better than younger adults on PM tasks in naturalistic-settings is minimising distractions (Phillips, Henry, & Martin, 2008), particularly if there is a short delay between when a cue to execute a PM task is presented and the opportunity to fulfil the intention (Einstein, McDaniel, Manzi, Cochran, & Baker, 2000). Supporting this interpretation, Einstein et al. (2000) found in a laboratory study that older adults in a non-divided attention condition showed equivalent performance after a delay as younger adults whose attention was divided. Furthermore, in naturalistic-settings older adults compared to younger adults are perhaps more likely, given their life experience, to recognise the potential for distractions between the forming of an intention and its execution, and thus find ways to reduce the delay for when a PM task can be actioned (McDaniel, Einstein, Stout, & Morgan, 2003).

1.4 Theoretical developments

1.4.1 The test-wait-test-exit model

After these early theoretical beginnings, theorists of PM began to focus on the importance of controlled attention, which appears to decline with age (Moscovitch, 1982), in the successful performance of PM tasks. The *test-wait-test-exit* model (Harris & Wilkins, 1982), in which people repeatedly and strategically shift their attention from whatever activity they are engaged in to check whether it is time to execute an intention, clearly has implications for older adults' likely performance on PM tasks. The strategic and cognitively demanding aspect of this theory is brought out in the findings of Harris and Wilkins (1982) that clock checking (i.e., the number of recorded times that a researcher observers a participant turn to look at the clock in a time-cued test of PM) tends to increase as the critical time for executing an intention approaches. To illustrate this, with an example from daily life that many older adults might face, consider the case of checking whether a chicken is cooked. The “tests” become more frequent when the chicken in the oven is close to being “just right”, that is, the window or “critical period” when the chicken has just cooked, and when delays beyond a particular duration will have the consequence of a burnt offering. The pattern of (typically) clock checking in younger adults tends to form a J-shaped pattern (Harris & Wilkins, 1982; Marsh, Hicks, & Cook, 2008); i.e., some clock checks at the beginning—perhaps to calibrate an internal clock—and then almost no clock checks until the time to execute the task approaches, in which phase checking becomes more and more frequent. Older adults tend not to show the same J-shaped pattern of clock checking (Einstein, McDaniel, Richardson, Guynn, & Cunfer, 1995; Park, Hertzog, Kidder, & Morrell, 1997) suggesting absorption in the distractor task or difficulty in strategic clock monitoring (cf. d'Ydewalle et al., 1999; Mioni & Stablum, 2014). And, indeed, one of the most influential early models of PM, the self-initiated process model (Craik, 1986), proposes that older adults should be particularly disadvantaged on PM tasks due to cognitive decline in controlled processing (see Chapter 2).

1.4.2 The self-initiated operations model

Although presented as a sketch in a chapter on ageing and memory, the self-initiated operations model presented by Craik (1986) is arguably the *locus classicus* in terms of early theorising about age-effects on PM. This model was developed from extrapolating from findings on retrospective memory paradigms in which age-effects increases as environmental support decreases. For example, free recall tests make high self-initiated processing demands; the participant needs to actively search their memory for the items to be recalled. In contrast, on a recognition test the target word is presented to the participant among distractors. The free recall case is low in environmental support, the participant is only given a cue by the researcher; whereas the recognition test is high in environmental support, the participant is not only given a cue to execute the task but also has the target word presented, which may lead to relatively reflexive memory processes. Craik (1986) argues that PM tasks are more akin to free recall (low in environmental support) compared to a recognition task (high in environmental support). However, according to Craik (1986) while PM is akin to free recall it is actually even *more* demanding of self-initiated operations than are free recall retrospective memory tasks. That is because PM tasks are assumed to offer not even the minimal environmental support of the researcher prompting the participant to recall the intention.

1.4.3 The Einstein and McDaniel PM laboratory paradigm

However, with the development of the classic Einstein and McDaniel laboratory paradigm of PM (Einstein & McDaniel, 1990), the view of all PM tasks being uniform in terms of minimal environmental support and demanding self-initiated cognitive operations needed revision. To understand the catalyst for this revision it is first necessary to present the procedure and typical content of what is now known as the “classic” or Einstein and McDaniel paradigm (Einstein & McDaniel, 1990). The classic paradigm is characterised by the following procedure (Einstein & McDaniel, 2005): 1) participants need to learn a distractor or ongoing task with practice trials (e.g., lexical decision task, such as classifying nouns as referring to animate or inanimate objects as quickly as possible); 2) participants are asked to perform a PM task (e.g., press a different key on computer if they see the word “tortoise”), which typically is introduced as of “secondary interest” (Einstein & McDaniel, 1990, p. 719), embedded within the ongoing task presented previously; 3) a filler task, i.e., some other activity unrelated to the practiced ongoing task, is given to prevent continuous rehearsal of the PM task; 4) the ongoing task is reintroduced (e.g., participants are given the lexical decision task instructions again) with no mention of the PM task; and 5) the target word (e.g., “tortoise”) is presented, typically 1–3 times, within the ongoing task. The first thing to be noted about the classic paradigm is that it (primarily) uses an event-based target cue for retrieval of the PM task to be executed (e.g., press key X when “tortoise” is presented). Having an event-based cue is indicative of at least some environmental support. Second, the target cue is *bivalent,* which means that it affords a response that could be consistent with either the ongoing task (classifying nouns as animate or inanimate) or the PM task (Einstein & McDaniel, 2005).

In a seminal series of experiments on normal ageing and PM using variations on the above procedure (e.g., a condition permitting improvising an external aid versus a no external aid condition), Einstein and McDaniel (1990) found no significant age differences between younger and older adults (all with above average intelligence) in PM for both external and no external aid conditions. In discussing how this finding diverged from what would be predicted by Craik (1986), Einstein and McDaniel (1990) identified two possibilities:

> The failure to find age differences in the no-aid condition suggests one of two possibilities. Either memory deficits in the elderly are not entirely due to problems in self-initiated retrieval, or certain kinds of prospective memory situations incorporate or have retrieval cues embedded in them. (p. 721)

In a second experiment Einstein and McDaniel (1990) manipulated the distinctiveness of the cue (i.e., familiar words versus unfamiliar words) relative to the stimuli in the ongoing task (i.e., lists of mostly familiar words) to investigate the possibility of spontaneous retrieval processes occurring when the target word was encountered. The improvement in PM performance for both young and older adults was approximately three fold. When using familiar word targets (i.e., non-distinctive cues), young and older adults' proportion correct was .28 and .36, respectively. In contrast, for the unfamiliar word targets (i.e., distinctive cues) young and older adults' proportion correct was .83 and .94, respectively. Age did not interact with familiarity condition and there were no main effects for age on PM proportion correct. Not surprisingly, given that younger adults were consistently superior to older adults on the diverse retrospective measures used as filler tasks in these experiments, retrospective memory was not significantly associated with PM performance.

It may appear from the foregoing that the distinction between event and time-based PM tasks had been established prior to this landmark paper by Einstein and McDaniel (1990). However, this is not the case. In the general discussion of the findings of the experiments in this paper, Einstein and McDaniel (1990, p. 724) draw for the first time the explicit distinction between prospective memory tasks involving either an event-based PM task, "the to-be-performed action is to be done when a certain external event occurs" (the example they give is passing on a message when receiving a phone call), or time-based PM task, "subjects must remember to perform some action after a period of time has elapsed or at a certain time" (the example they give is attending a 4pm meeting). Based on their findings they propose that event-based PM tasks are not affected by ageing, at least if the retrospective memory component is simple (Einstein, Holland, McDaniel, & Guynn, 1992), while proposing that time-based tasks appear to be particularly reliant on self-initiated processes, and thus should show normal age-effects similar to those found in free recall retrospective memory tests (Einstein et al., 1995).

However, as the current research project seeks to argue and demonstrate empirically, time-based tasks, like event-based tasks, are *not* necessarily dependent on self-initiated processes. This is particularly so in naturalistic-settings where a constellation of time related cues in the environment may lead to the spontaneous retrieval of a delayed intention, such as attending a meeting at 4pm. This insight has been occasionally recognised in the age-PM literature but frequently seems to slip researchers' minds when they treat time-based tasks as essentially homogenous. Rabbitt (1996) deserves acknowledgement for recognising this point early on in the PM literature when he wrote:

> Craik's (1986) suggestion that, in everyday life, time-based prospective memory is unique because it receives minimal environmental support misses the point that knowledge of the passage of time is impossible without environmental support. This is not a trivial semantic quibble because, in everyday life, the changes that people continually experience or make to their environments, to themselves, or to their locations in space provide directly perceptible sequences of "self-updating" cues that mark the passage of time and also provide unambiguous agendas to schedule their actions. (p. 241)

Thus, PM tasks may prove more or less demanding of controlled attentional processes, and, perhaps, on occasion be executed relying on spontaneous processes unaffected by age related cognitive changes. This notion suggests multiple processes may be involved in the fulfilling of PM tasks.

1.4.4 The multiprocess framework

In the fruitful decade of research on ageing and PM following the establishment of the Einstein and McDaniel (1990) paradigm, a pattern of contradictory results emerged in terms of the association of ageing to PM performance on both time (d'Ydewalle et al., 1999) and (primarily) event-based tasks (Einstein et al., 1992). To make sense of this pattern of results, and to further the research on age-effects and PM task characteristics, a new model of age-effects on PM was proposed, the *multiprocess framework* (McDaniel & Einstein, 2000). The multiprocess framework proposed that two fundamentally different retrieval processes are involved when a cue or target to execute a PM task is encountered. The first process occurs when the cue is distinctive or related to the cognitive processing involved in the ongoing task, and is characterised by relatively spontaneous retrieval; which, like recognition memory, is relatively unaffected by the cognitive ageing process (Kliegel, Jäger, & Phillips, 2008). The second process is the one previously identified by Craik (1986) involving self-initiated operations, in particular, voluntary, strategic monitoring processes. McDaniel and Einstein (2000) also noted that it is possible for the one PM task to involve both processes, such as when the task to be executed is complex or the association between the target cue and the content of the intention to be executed is weak. For example, imagine going to a shopping center to have a coffee with a friend. When you enter the center you may spontaneously recall a previously formed intention to by milk from the supermarket. However, if the intention was to buy another item, not so readily associated with seeing the shopping center (e.g., envelopes), then a two-step process would occur. First, a spontaneous retrieval that “something” needed to be purchased, and second an effortful, strategic process of searching for the “what” (i.e., envelopes) needed to be purchased. The claim of the multiprocess framework is that for simple delayed intentions (i.e., involving a distinctive cue or natural association between cue and action) then relatively spontaneous or reflexive associative processes will be recruited and age related PM deficits will not occur (Einstein & McDaniel, 1990).

1.4.5 The PAM model

An alternative theory of PM, which also has applicability to cognitive age effects, is the *preparatory attention and memory processes* (PAM) theory (Smith & Bayen, 2004). The PAM model proposes that attentional processes demanding of cognitive resources are always involved in any PM task, and that thus there is always a “cost”, measured by either decreased performance in an ongoing task (i.e., whatever activity a person is engaged in during the delay between the formation of the intention and the appearance of the cue signaling that the PM task is due to be executed) or decreased performance on the PM task itself. In other words, according to the PAM theory, there is always a trade-off, perhaps more noticeably so for older adults, between remembering to execute a future intention successfully and performing well on whatever other tasks are at hand (Smith & Bayen, 2004).

1.4.6 Summary

This synoptic overview of the early and more recent theories of PM that have been used to explain patterns of age differences in PM performance, indicates that a central concern expressed in seminal works in the PM literature (e.g., Craik, 1986; Einstein & McDaniel, 1990; Maylor, 1990; Moscovitch, 1982; Rendell & Craik, 2000; Rendell & Thomson, 1999) is how cognitive changes correlated with ageing interact with the relative demandingness of a PM task; and in particular, what environmental support is available to help retrieve the intention and execute it at the appropriate time (Rendell & Craik, 2000). The following section critically considers some of the research to date that has both shaped and been shaped by the theoretical developments outlined above (e.g., the multiprocess framework). This empirical research provides a complementary rationale, to that given by the theory, for the undertaking of the current research project.

1.5 What we currently know about cognitive ageing

There is variability both between and within older individuals in terms of the decline of different cognitive processes, primarily non-verbal fluid abilities, such as perceptual processing speed, mental set shifting, inhibition, and aspects of working memory, that are theorised to underpin PM performance (Kliegel, Martin, McDaniel, & Einstein, 2002; Salthouse, Berish, & Siedlecki, 2004; Schnitzspahn, Stahl, Zeintl, Kaller, & Kliegel, 2013; Zeintl, Kliegel, & Hofer, 2007; Zuber, Kliegel, & Ihle, 2016). There is some evidence that older adults who have had more years of education may show reduced declines due to *cognitive reserve* (Cherry & Lecompte, 1999; Stern, 2012), a term that refers to a capacity to sustain a particular level of cognitive performance despite structural changes in the brain associated with ageing (Park & Reuter-Lorenz, 2009). There is also longitudinal empirical evidence that most cognitive decline occurs after 60 years of age (Salthouse, 2016). Furthermore, this decline does not appear to be uniform across cognitive domains, at least for the young-old, as their crystalized intelligence and verbal IQ are frequently superior to that of young adults (Hofer & Alwin, 2008). On the other hand, there appears to be robust evidence that the ability to shift attention between tasks and control attention (e.g., inhibiting distractions or avoiding mind wandering; Maillet & Schacter, 2016) does decline with age. To maintain more than one goal (as is the case for PM tasks, where performing an ongoing task is also a goal), it is often necessary to be able to strategically shift attention (Schnitzspahn et al., 2013). Redirecting attention from one task or mental set to another (e.g., strategically checking the time, or changing mental sets when a target cue appears as part of the ongoing task) is a ubiquitous phenomenon that is no less relevant for older adults than it is for young adults. In the case of event-based tasks this involves *seeing the stimuli as a PM cue or target, and responding appropriately*, rather than simply part of another unrelated, ongoing task.

Immediately subsequent to this *seeing as* or recognising of a PM cue, it is frequently necessary to be able to inhibit an ongoing prepotent response in order to execute a PM task (Kliegel et al., 2002).

There is some debate as to whether higher order cognitive processes (i.e., complex cognitive processes such as working memory and PM) decline as a function of lower order cognitive processes (e.g., processing speed; Salthouse, 1991, 1996; Verhaeghen, 2014), but in general, many researchers have subscribed to the frontal lobe hypothesis (Volle, Gonen-Yaacovi, de Lacy Costello, Gilbert, & Burgess, 2011; West, 1996), whereby cognitive changes associated with age are largely a function of declines in attentional processes controlled by the frontal cortex. In particular, researchers have repeatedly noted the resemblance between older adults' declining performance on measures of executive functioning and the dysexecutive syndrome shown by patients who have injuries to the prefrontal cortex (Hofer & Alwin, 2008; West, 1996). However, there is a well-known "impurity" problem with measures of executive functioning (i.e., they measure more than one construct; see Miyake et al., 2000), and the general or overarching construct of "executive functioning" is somewhat vague. Thus it is not exactly clear what is meant when it is claimed that executive functioning (as a general construct) mediates age-effects on PM. One response has been to "fractionate" the central executive (Baddeley, 1996) and take a fine grained approach to which facets of executive functioning mediate age-effects on a range of PM tasks. The division of executive functioning into facets in ageing research has largely followed the latent variable analysis of Miyake (2000), who factor analysed a range of commonly used executive functioning measures (with young adults it should be noted). Miyake et al. (2000) identified three distinct facets of executive functioning: shifting (the ability to switch between rules of responding), updating (the ability to manipulate the contents of working memory), and inhibition (the ability to avoid distraction or pre-potent responses when required). Chapter 5 of the current research project continues this work (Schnitzspahn et al., 2013; Zuber et al., 2016), investigating how shifting, updating, and inhibition, mediates age-effects on a wide range of PM measures.

1.6 What we currently know about ageing and prospective memory

There has been some debate as to whether PM necessarily declines with age (Uttl, 2008), which has largely been driven by discrepant age-related findings across laboratory and naturalistic-settings (Henry et al., 2004). The key to resolving the apparent age-PM paradox appears to be in considering both the characteristics of the PM task (and competing ongoing task) and how these interact with individual differences in executive functioning or other cognitive processes. In general, the more “demanding” the ongoing activity or PM task to be executed, the more apparent is age-related decline in PM performance (Einstein et al., 1992; McDaniel & Einstein, 2008; Mullet et al., 2013). One study supporting this conceptualisation found that young adults who were given a complex ongoing task to perform, and had a short delay between presentation of the target cue and the opportunity to execute the PM task, performed at the same level as older adults who had only a simple ongoing task but the same delay-execute PM task (Einstein et al., 2000). However, older adults who were in the same complex ongoing task condition showed substantially poorer PM performance than the young adults in the same condition. This finding suggests that both young and older adults have limited attentional resources, only the latter more so. In other words, when younger adults’ attention is duly taxed or divided on an ongoing task both before and during the appearance of the cue to execute the PM task, then they will tend to perform at an equivalent level to older adults whose attentional resources are not so taxed.

In two influential meta-analyses of ageing and PM, both Henry et al. (2004) and Kliegel, Jäger, et al. (2008) found that age-effects were most evident on high demand tasks, such as when a cue to execute a task was not related to the ongoing task (non-focal) or not distinctive relative to the ongoing task. An example of one of these high demand tasks is where the ongoing task is writing down the name of famous faces appearing on slides and the PM task is simply to circle the number of any slide which has the picture of a person wearing glasses (Maylor, 1996; Rendell, McDaniel, Forbes, & Einstein, 2007). This is an example of a non-focal task, as the processing required for the ongoing task (recognising and writing down the name of the famous face) does not necessarily involve attention being directed to whether or not the famous face is wearing glasses. This is an important example of a high demand task as it draws attention to the fact that a non-focal cue does not need to be peripheral relative to an individual's visual field, but simply unrelated to the cognitive demands of another task.

Few studies have investigated high demand PM tasks in a naturalistic-setting with older adults. In one of the few studies, Niedźwieńska and Barzykowski (2012) manipulated cue focality by asking young and older adults who reported regularly watching the news to contact the researcher when the map of Poland appeared in the weather report (focal cue, relatively low demand), or at the first mention of a Polish politician's name in the news (non-focal cue, relatively high demand). The typical paradoxical pattern of older adults performing better than young adults emerged in the low demand condition, but in the high demand condition there was no difference between the two age groups. Rendell and Craik (2000) varied PM task demands in a naturalistic-setting and found that older adults outperformed young adults on time-of-day and event based tasks. However, in the same study, for a time-interval task (with 30 and 60 minute intervals) there was no difference between the old-old and young adults, while the young-old adults performed better than both of these groups. In Chapter 4 of the present research project further work is carried out along the lines of Rendell and Craik (2000) using a range of PM tasks in a naturalistic-setting, while in Chapter 5 how older adults' performance on executive functioning measures predicts their PM performance is also investigated.

The current research project (Chapter 4) argues that the apparent age-PM paradox may partly be a function of *not* using conceptually similar PM tasks across settings. However, the results of this investigation suggest there may be some fundamental challenges in creating conceptually similar tasks across settings, such that even apparently "low demand" tasks for older adults in the laboratory can seemingly become high demand when embedded within multiple other different PM task types in the one paradigm (Virtual Week in this case).

1.7 Research paradigms in prospective memory and ageing research

Various paradigms have been used in investigating age-effects on PM. However, one clear distinction among these paradigms corresponds to whether participants interact with the paradigms and execute the PM tasks in a laboratory or naturalistic-setting. Laboratory paradigms have largely been marked by high levels of experimenter control, while naturalistic paradigms have included more familiar PM tasks (e.g., mailing postcards) but have not had the same level of experimenter control. The development of the influential Einstein and McDaniel (1990) paradigm, with typically 1–3 PM target cues within a demanding ongoing task (e.g., lexical decision making), was developed in response to the low levels of experimenter control in earlier, mostly naturalistic-setting paradigms used to investigate age-effects on PM. Examples of PM tasks embedded within these earlier paradigms include sending postcards on specified days (Patton & Meit, 1993) or making a phone call to the experimenter at an agreed upon specific or general time (Maylor, 1990). Early laboratory experiments of PM included tasks such as reminding the experimenter to put a phone back on the hook at the end of a study (Kvavilashvili, 1987) or asking for a red pen when the experimenter gave a particular, indirect cue (Dobbs & Rule, 1987).

A paradigm that combines aspects of both the control of the laboratory and familiar PM tasks of daily life (which may have given older adults an advantage in naturalistic-settings) is Virtual Week (Rendell & Craik, 2000; Rendell & Henry, 2009). In the Virtual Week laboratory paradigm, participants are engaged in a board game ongoing task, in which they imaginatively engage in a series of daily life decisions, PM tasks, and commonplace events as they move a token around the board. The hypothetical daily life events, “time of day”, decision making options, and PM tasks are all controlled by the experimenter. (A more detailed account of Virtual Week is presented in the Methodology, Chapter 3).

Naturalistic paradigms for studying age-effects on PM are currently becoming more common and sophisticated in terms of experimenter control (Bailey, Henry, Rendell, Phillips, & Kliegel, 2010; Niedźwieńska & Barzykowski, 2012; Rose et al., 2015; Schnitzspahn, Ihle, Henry, Rendell, & Kliegel, 2011). These new paradigms have some resemblance to the earliest paradigms previously mentioned (e.g., sending postcards) but the specific PM tasks are more contemporary (e.g., sending text messages from mobile phone). They also involve cues that the experimenter can control, such as a cue embedded within a survey that participants need to complete on a smartphone app (Bailey et al., 2010). The current research project continues this trend with a relatively new naturalistic PM paradigm, MEMO (Randall, 2016), that has parallel PM tasks to those administered in Virtual Week.

1.8 Investigating age-effects on diverse high and low demand PM tasks

There have been relatively few attempts to manipulate the cognitive demands (and thus potential age effects) of PM tasks across settings (Bailey et al., 2010; Niedźwieńska & Barzykowski, 2012; Rendell & Craik, 2000; Schnitzspahn et al., 2011). It is important to pursue this research program in order to 1) probe the age-PM paradox, whereby older adults generally show pronounced PM deficits on high demand PM tasks, relative to young adults in laboratory-settings, while in naturalistic-settings this pattern is reversed; and 2) to continue investigating the general theoretical proposition, and extend the empirical (mostly laboratory) research, on the mediating role of facets of executive functioning ability on a diverse range of putatively high and low demand PM tasks. Thus, the key aims in this research project were to investigate how PM task types modulate the effect of age in laboratory and naturalistic-settings and potentially shed light on mechanisms driving the age-PM paradox; and two, to investigate whether facets of executive functioning ability mediate age-effects on a diverse range of PM tasks.

1.9 Next steps

In this general introductory chapter, the ground has been laid for the empirical work described in Chapter 4 (an investigation of parallel tasks across settings to resolve the age-PM paradox) and Chapter 5 (an investigation of the associations between executive functioning abilities and how these might account for age-effects on PM tasks, particularly high demand tasks). The following literature review (Chapter 2) provides a detailed account of the research on cognitive ageing and PM, which builds on this introductory chapter. Chapter 3 provides an extended description of the methodology utilised and is vital in understanding how this research project fits in with other research on cognitive ageing and PM (particularly in regards to measuring executive functioning and other typically associated cognitive processes with PM, as well as the diverse PM measures used in this thesis).

The next steps in the age-PM field are to actively manipulate PM task characteristics in a naturalistic-setting to improve our understanding of how age-effects interact with diverse PM task characteristics that are encountered in daily life. The findings to date, that older adults are doing fine in naturalistic-setting but poorly in the laboratory, may invite the question: why should we be concerned about age-effects on PM? The answer to this is that we do not yet know enough about the diverse types of PM task characteristics faced by older adults in naturalistic-settings. Thus to better understand age-related PM lapses it is necessary to investigate a conceptually similar array of PM task characteristics in both laboratory and naturalistic-settings.

Chapter 2 Review Paper of Prospective Memory and Cognitive Ageing

Preamble

This chapter was submitted as an article for publication on the 22nd June, 2018 to *Oxford Research Encyclopedia of Psychology*. It was accepted with revisions on 2nd September, 2018 and published online 28th February, 2019. It is presented in the format in which it was submitted prior to being accepted (i.e., after responses were made to reviewers' feedback). However, additional footnotes have been added as part of the thesis process. This is where examiners of the thesis have identified or raised issues. Each of these footnotes clearly states that it has been added post production, as part of the thesis process, to address issues raised by thesis examiners, and does not appear in the published chapter. For evidence of this chapter as an accepted publication see Appendix C–2. The reference for this publication is:

Haines, S. J., Shelton, J. T., Henry, J. D., Terrett, G., Vorwerk, T., & Rendell, P. G. (2019). Prospective memory and cognitive aging. *Oxford Research Encyclopedia of Psychology*: Oxford University Press. doi: 10.1093/acrefore/9780190236557.013.381

Statement of Contribution for Accepted Publication:

Prospective Memory and Cognitive Aging

Submitted to *Oxford Research Encyclopedia of Psychology* on 22^{nd} June, 2018, accepted with revisions on 2^{nd} September, 2018, published online 28^{th} February, 2019

I acknowledge that my contribution to the paper is 50%

Name: Simon Haines

Signature:

[redacted]

I acknowledge that my contribution to the paper is 10%

Name: Jill Talley Shelton

[redacted]

I acknowledge that my contribution to the paper is 10%

Name: Julie D. Henry

Signature:

[redacted]

I acknowledge that my contribution to the paper is 10%

Name: Gill Terrett

Signature:

[redacted]

I acknowledge that my contribution to the paper is 10%

Name: Thomas Vorwerk

Signature:

[redacted]

I acknowledge that my contribution to the paper is 10%

Name: Peter G. Rendell

Signature:

[redacted]

Prospective Memory and Cognitive Aging

Summary: Tasks that involve remembering to carry out future intentions (such as remembering to attend an appointment), and the cognitive processes that enable the completion of such tasks (such as planning), are referred to as prospective memory (PM). PM is important for promoting quality of life across many different domains. For instance, failures in remembering to meet social commitments are linked to social isolation, whereas failures in remembering to fulfil occupational goals are linked to poorer vocational outcomes. Declines in PM functioning are of particular concern for older adults because of the strong links between PM and functional capacity. The relationship between age and PM appears to be complex, dependent on many different factors. While some aspects of PM appear to hold up relatively well in late adulthood, others appear to show consistent age-related decline. Variability in age differences appears to partially reflect the fact that there are diverse types of PM tasks, which impose demands on a range of cognitive processes that are differentially affected by aging. Specifically, the level and type of environmental support associated with different PM task types appears to be a meaningful determinant of age-related effects. Given the worldwide changing age demographics, the interest in age-related effects on PM will likely intensify, and a primary focus will be how to optimize and maintain PM capacity for this population. This is already reflected in the increasing research on interventions focused on enhancing PM capacity in late adulthood, and points to important future directions in this area of study.

Keywords: prospective memory; event-based; time-based; cognitive aging; individual differences; spontaneous retrieval; strategic monitoring; environmental support

Introduction

When people are asked to recall a time when they had a memory lapse they may produce a story about some information that they knew but could not remember when prompted (e.g., the answer to an exam question) or about something they unintentionally omitted to do (e.g., attend a meeting). The former lapse involves *retrospective memory,* the ability to recall either a unit of information or an event that was previously encoded and stored in long term memory. The latter lapse involves a failure of *prospective memory* (PM), which subsumes many of the processes involved in retrospective memory, but necessarily also involves the encoding of a future intention (McDaniel & Einstein, 2007) and *self-initiated processes* (Craik, 1986) to fulfil that intention. Self-initiated processes are those that depend on the agent’s own initiative, e.g., keeping an eye on the time or setting up a reminder, and may involve, to some extent, executive functions, such as inhibition, updating, and task switching, which are known to decline with age. An essential feature of PM tasks is that there is a gap between forming an intention and the opportunity to execute it, where the length of gap is not important as it can range from minutes to days. The critical feature is that during the gap there are other activities that need to be attended to and these are referred to in the PM literature as the ongoing task. As well as the variable nature of the ongoing task, PM tasks may also differ in environmental support, that is, the concurrent or external cues available in the environment (e.g., a familiar television program that could convey time of day or a note on the fridge) that might act as a signal to the agent to recall and execute a delayed intention. Environmental support is a key concept in understanding age-related effects on PM, as they are[3] often absent in the laboratory, and plentiful in some daily life contexts, but sparse in others. Given that PM failures dominate memory complaints, with studies reporting 50-80% of everyday memory problems are PM errors (Kliegel & Martin, 2003),

[3]Issue raised post publication: *one examiner identified a typo*: “they are” should be “it is”.

understanding this ability is clearly important, especially in relation to changes in this capacity as we age. The following sections show how models of PM and aging have developed as a function of considering environmental support and PM task types.

Models of PM and Aging

Self-initiated Operations

Early PM theorizing grew out of consideration of the different patterns of performance on retrospective memory paradigms in the lab (Craik, 1986), such that retrospective memory was good when some conditions were present and poor for others. Particularly influential was Craik's (1986) self-initiated model of PM processing, in which a central tenet was that for memory tasks high in environmental support (e.g., recognition) older adults' performance is less negatively affected relative to tasks that are low in environmental support (e.g., free recall, PM). However, the diversity of PM tasks suggests that multiple cognitive processes underlie successful prospective remembering (not just self-initiated ones) and these are differentially affected by the aging process.

The Multi-Process Framework

The multiprocess framework (McDaniel & Einstein, 2000, 2007; McDaniel, Umanath, Einstein, & Waldum, 2015) proposes multiple processes are involved in PM. The default process—particularly for older adults—is to rely on relatively automatic, reflexive processes whereby cues in the environment spontaneously elicit the PM task to be performed. One benefit of relying on more automatic processes to execute PM tasks is that more cognitive resources are available to support other ongoing activities (e.g., engaging in a conversation with a friend). Some ongoing or PM tasks are, however, more cognitively taxing than others and require greater attentional control. For example, relying on automatic processes would be risky when there is a tenuous association between an environmental cue and the PM target (e.g., buy shoe polish at supermarket) or when a PM task is critical (e.g., buy medicine at pharmacy). In such cases, the successful completion of PM tasks may hinder performance on other ongoing tasks (e.g., not attending well to a conversation with a friend). Consistent with predictions from the multi-process framework, PM tasks well suited to relatively automatic processes have revealed only minimal age effects; whereas those involving more controlled processing typically show pronounced, negative age-effects for performance of the PM and/or ongoing task (Kliegel, Jäger, et al. (2008).

The Dynamic Multi-Process View (DPMV)

Extending the multiprocess framework, the dynamic multiprocess framework (Scullin, McDaniel, & Shelton, 2013) contends that people flexibly switch between bottom-up (e.g., spontaneous retrieval) and top-down processes (e.g., strategic monitoring) throughout intention encoding, maintenance, retrieval, and de-activation stages to support prospective remembering. The dynamic multiprocess framework emphasizes the importance of contextual and individual difference factors in modulating the interplay between bottom-up and top-down processes supporting prospective remembering. This notion of flexibility is based on the observation that, unlike most lab studies, everyday PM tasks are often complex with longer, sometimes unpredictable, delays and embedded in diverse conditions, including high cognitive load, fatigue, and variable environmental support.

Thus, it is hypothesized that PM age-related effects may not simply be a function of insufficient cognitive resources. Rather, age-related effects may also stem from less flexibility in switching between automatic and strategic processes in response to changing environmental demands, as well as a variety of individual difference factors (e.g., cognitive ability, motivation, daily routines). This raises the question of the role of meta-memory—individuals' awareness of their own memory capacity and fallibility in different situations. For example, if an older adult knows that they tend to forget to execute their future intentions while balancing demands of a busy day, they could plan to have greater environmental support throughout those days (e.g., strategically placed reminders) to increase the probability of spontaneously retrieving intentions at an opportune moment.

Preparatory Attention and Memory Processes (PAM) Model

In contrast to the DMPV, the PAM model (Smith, 2003) posits that all PM tasks incur a measurable behavioral cost to ongoing activities. Given that cognitive processing resources—such as attentional and working memory capacity—decline with age (Park et al., 1997), these costs are likely to be more pronounced for older compared to younger adults. In support of the PAM model, Smith (2003) found that performance on a standard laboratory ongoing task was compromised when a PM task was subsequently embedded. In this model there are no cost free, automatic processes involved in PM at any age or for any ongoing task (Bisiacchi, Tarantino, & Ciccola, 2008; Smith, Horn, & Bayen, 2012); instead a continuous process of monitoring for the PM cue, followed by effortful retrieval of the PM content (cf. Marsh, Hicks, & Landau, 1998), is always present.[4]

---

[4]Issue raised post publication: *it was commented by one examiner that there is a "relatively new" theory of PM not specifically related to ageing that could be mentioned*. This is the Delay Theory (Heathcote, Loft, & Remington, 2015; Strickland, Loft, Remington, & Heathcote, 2018) of event-based PM, which provides a somewhat different interpretation of the pattern of PM "costs" found on ongoing task performance. In essence this theory proposes that rather than one pool of divided attentional resources accounting for event-based PM costs, there are variable "competing" evidence accumulation processes involved. These processes compete for the decision to be made in response to stimuli that are relevant to the ongoing task and potentially also serve as the PM target cue. The threshold for this decision making can be mathematically modelled using sophisticated statistical techniques, such as linear ballistic accumulator modelling. As this is both a relatively new theory and not specifically targeted at understanding age-related effects on PM tasks it is not presented and discussed in the main text. However, it might be speculated that the biases for "decision making" processes may change with ageing, similar to the proactive and reactive response biases proposed by Braver and colleagues (Braver & Barch, 2002; Braver et al., 2001; Braver, Reynolds, & Donaldson, 2003; Braver, Satpute, Rush, Racine, & Barch, 2005).

Taken together, theories of PM agree that retrieval of future intentions *may* require top-down, self-initiated processes that consume cognitive resources. In such situations, all theories would predict older adults to perform more poorly than younger adults. The multiprocess, or dynamic multiprocess framework, uniquely predicts that there are situations in which the retrieval of future intentions can be elicited through a more reflexive, bottom-up process that is preserved with age. Throughout the chapter, the empirical findings in the PM and aging literature are evaluated through the lens of these different theoretical perspectives.

Approaches to Studying PM and Aging

A variety of approaches have been used to examine PM across the lifespan; the cognitive and motivational factors that moderate PM with aging; and how contextual factors, such as environmental support influence PM. Historically, PM has been divided into event- and time-based categories. Event-based tasks reflect instances in which PM task execution is marked by a particular event, such as remembering to buy apples when you drive past the market. Time-based tasks can be fruitfully distinguished into two subtypes: *time-of-day* (e.g., attend a doctor's appointment at 11am) or *time-interval* tasks (e.g., turn the oven off in 10 minutes) marking appropriate time for PM execution. Laboratory PM studies are dominated by event-based and time-interval tasks, requiring monitoring intervals of up to 20 minutes; while naturalistic PM studies are dominated by appointment type time-based tasks, requiring monitoring time-of-day. The key difference between event- and time-based tasks is the type of cue used to signal PM execution. Within these broad categories, a key distinction within event-based tasks is whether the cue is related or unrelated to the ongoing task (referred to in the PM literature as *focal* vs. *non-focal* cue, respectively). On the other hand, for time-based tasks an often overlooked distinction exists between time-interval and time-of-day tasks. Namely, the former are low in environmental support due to typically few or no changes in the environment indicating the passing of the relevant duration of seconds or minutes signaling execution; while the latter are high in environmental support due to many changes in the environment indicating time-of-day that could signal the time for execution. One reason this latter distinction may have been overlooked is due to the typical PM tasks used in laboratory (time-interval) and naturalistic (time-of-day) settings that may have obscured important PM age-related effects. Nevertheless, PM and aging researchers have systematically manipulated the level of environmental support by exploiting these distinctions (e.g., focal vs. non-focal) in their study designs to identify age-related effects in both laboratory and naturalistic-settings.

Lab and Naturalistic PM Tasks

Early studies on PM involved participants completing experimenter given time-based tasks in a naturalistic-setting, such as mailing postcards on designated days or calling the experimenter at designated times (Meacham & Singer, 1977; Moscovitch, 1982). Later studies continued to use phone calls as a PM task, which provided a relatively accurate measure of PM and a convenient way of classifying early or tardy responses (Aberle, Rendell, Rose, McDaniel, & Kliegel, 2010; Kvavilashvili & Fisher, 2007; Maylor, 1990). The first large scale (n = 320) naturalistic-setting PM study with older adults was by Maylor (1990). This study used regular phone calls either at an exact time or between two times providing a four hour window of opportunity. It was found that participants performed better on the exact time task and that this correlated with self-reports of using conjunction cues (i.e., associating required time-of-day for telephone call with an event that typically occurred at that time). Maylor’s (1990) use of time-of-day, time-based PM tasks is representative of the majority of research conducted in naturalistic-settings, and also occurred at the same time as one of the first systematic, non-clinical laboratory measures of PM was developed and reported in a seminal study by Einstein and McDaniel (1990).

The classic event-based lab paradigm, also known as the Einstein and McDaniel (1990) paradigm, was developed to examine age-related effects on PM relative to retrospective memory. In this paradigm a small number of PM target words occurred infrequently during an ongoing short-term memory task with lists of words. Performance was operationalized as the proportion of times participants made the required response to the PM target. Notably, Einstein and McDaniel (1990) did not find age deficits in event-based PM which led to the optimistic speculation that PM "seems to be an exciting exception to typically found age-related decrements in memory" (p.724).

The Einstein and McDaniel (1990) paradigm has been further developed to identify factors supporting retrieval of future intentions. To capture the amount of attentional resources being devoted to the PM task, speed and accuracy in the ongoing task is compared between conditions when a PM demand is present relative to when there is no PM demand (Marsh, Hicks, Cook, Hansen, & Pallos, 2003; Smith, 2003; see also section above on PAM model)[5]. Importantly, the level of environmental or contextual cues available to support PM during the ongoing task modulates the cognitive processes needed for successful retrieval, which has important implications for observed age differences (McDaniel & Einstein, 2007).

It is noteworthy that while evaluating PM in real-world contexts is important, the lab-based approach offers the complementary experimental control needed to isolate the cognitive processes supporting retrieval of future intentions. In lab-based studies, to address the theoretical question of how much attention is devoted to fulfilling short-time interval time-based tasks,

---

[5]Issue raised post publication: *it was commented by one examiner that the following reference could be cited for the neurological underpinnings of these attentional resources*, this reference is Burgess, Quayle, and Frith (2001).

researchers have recorded the frequency and rate of clock-checking prior to the time for intention execution (Ceci & Bronfenbrenner, 1985). The frequency and trajectory of clock-checking provides an index of attentional allocation to the PM task. In contrast to the early phases of the life span where a J-shaped pattern of clock checking emerges (i.e., more checking as the critical time approaches)—with initial frequency of checking suggesting an internal clock has been calibrated—older adults tend to demonstrate less strategic clock-checking patterns when there are constraints on the number of clock checks allowed (Mioni & Stablum, 2014).

A key lab paradigm that explicitly aimed to simulate the contextual support found in daily life is Virtual Week (VW; Rendell & Craik, 2000; Rendell & Henry, 2009). VW is a computer-based task that simulates daily life activities using a board game format (see Figure 1). Participants move around the board with the roll of a die; each circuit around the board represents one virtual day. At the center of the board game, there is a virtual clock; as participants move the token around the board, the virtual time changes with every two squares representing 15 minutes. There is also a stop clock at top of screen showing the real time elapsed since start of the virtual day. The two clocks in VW allow for "time-of-day" (e.g., "at 1 pm get haircut") and time-interval tasks (e.g., "check lung capacity at 2:00 and 4:00 minutes on stop clock").

Figure 1. Virtual Week (Rendell & Craik, 2000) computerized version board game user interface.

While moving around the VW board, participants must pick up “Event cards” showing activities relevant to the virtual time-of-day (e.g., eating meals and going shopping) and choose a response from options presented on the card regarding those activities (e.g., select what to eat or buy). Participants also have to remember to perform additional tasks, which closely resemble PM tasks in daily life (e.g., buy gift while shopping), by clicking on the “Perform Task” button and selecting the appropriate task from a list that includes distractors. The time-based and event-based PM tasks can be *regular* (PM tasks recurring each virtual day; e.g., take antibiotics with breakfast and dinner; use asthma inhaler at 11 am and 9 pm) or *irregular* (one off PM tasks; e.g. “phone the bank at noon” and “pick up membership card when at swimming today”).

The first VW study by Rendell and Craik (2000; Study 1) compared young, young-old, and old-old adults on regular, irregular, and time-check tasks. They found that the young adults consistently outperformed the older adult groups with the exception of regular tasks in which there was no difference between young and young-old (both these groups outperformed old-old). Young-old and old-old showed comparable performance except on the irregular tasks, for which young-old outperformed old-old. These findings were contrasted in a follow up study by Rendell and Craik (2000; Study 2) which found superior performance by older adults compared to young adults on parallel tasks in a naturalistic-setting ('Actual Week'). However, a similar pattern within age-groups on task types emerged, e.g., poorer performance on time-check task than event-based tasks (see Age-PM Paradox section), showing that VW is partially indicative of the interaction of age with PM task types in a naturalistic-setting.

The VW paradigm has proven to be a fruitful approach to advancing research on aging and PM. For example, using a version of VW with emotionally valanced PM tasks, Rendell et al. (2011) found evidence for positivity enhancement—successfully performing more positive than neutral PM tasks—that was considerably larger for older than young adults. This was found only for event-based tasks; no negative enhancement or impairment (negative vs neutral) was found. VW has also been translated into many languages (e.g., Mioni, Stablum, Biernacki, & Rendell, 2015, reports the validation of the Italian version; Niedźwieńska, Rendell, Barzykowski, & Leszczyńska, 2016, reports the validation of Polish version) and used to show collaboration by long term older couples on PM tasks (Browning, Harris, Van Bergen, Barnier, & Rendell, 2018).

Age-PM Paradox

A robust finding in the PM and aging literature is that older adults perform poorly on laboratory PM tasks with high attention demands, but are superior on most naturalistic PM tasks, relative to younger adults—a pattern known as the age-PM paradox (Moscovitch, 1982; Rendell & Craik, 2000; Rendell & Thomson, 1999) . Although the age-PM paradox has typically been thought of in relation to extreme age group contrasts, an overlooked aspect of the paradox is the comparison of young-old with old-old. When comparing young-old and old-old there is often large age-related declines in laboratory PM studies but no age differences in naturalistic PM studies (Kliegel, Rendell, et al., 2008). In a meta-analysis of age-related effects on PM, Henry et al. (2004) found performance on time and event-based PM tasks were negatively correlated with age in the lab, but positively correlated in naturalistic-settings—the effect-sizes of these setting-dependent relationships were mirrored in magnitude (Phillips et al., 2008). The PM paradox is considered one of the most perplexing findings in the cognitive aging literature (Kliegel et al., 2016); motivating much research on possible factors driving this pattern (for an argument that this pattern arises through methodological artifacts, see Uttl, 2008). As detailed next, this research has unveiled a number of contributing factors, as well as boundary conditions.

Age Differences in Event-based PM Tasks across Settings

There is evidence to suggest that some event-based tasks may be relatively insensitive to the PM paradox, which appears to be primarily driven by the comparison of (disparate) time-based tasks across lab and naturalistic-settings (Kliegel, Rendell, et al., 2008; Randall, 2016). Kvavilashvili, Cockburn, and Kornbrot (2013) found no difference in event-based tasks between young and young-old on two event-based tasks in the laboratory (though both these groups were superior to old-old) and no age-related effect in a naturalistic-setting. Furthermore, on one event-based task (remembering to retrieve a wrist watch) young-old were superior to both young and old-old adults. Niedźwieńska and Barzykowski (2012) found no age-related effect in the lab on a focal event-based task, and no age-related effect on a non-focal event-based task in a naturalistic-setting (middle aged and older adults did however outperform young adults on a focal event-based task, while not differing from each other).

In addition, for event-based PM tasks in the lab, older adults do not always exhibit difficulties relative to their younger counterparts (cf., Dobbs & Rule, 1987). Specifically, for event based tasks with salient or focal cues older adults are only marginally negatively affected in terms of performance in formal test settings relative to younger adults (Henry et al., 2004;

Uttl, 2008). Instead, it appears to only be for more demanding event-based tasks (e.g., when the cue is less clearly associated with the target response-or is nonfocal), that older adults show a marked decrease in performance (Henry et al., 2004). Similarly, higher intra-individual variability in costs for ongoing task response time (i.e., sometimes longer and sometimes shorter delays in responding) have been related to increased age only in a non-focal cue condition (Ihle, Ghisletta, & Kliegel, 2016), suggesting lapses in controlled cognitive processes. An important caveat, though, is that in high cognitive load conditions in the lab, older adults have been found to exhibit equivalently reduced PM performance for both focal and non-focal tasks (Ballhausen, Schnitzspahn, Horn, and Kliegel (2017).

Age Differences in Time-based PM Tasks across Settings

The age-PM paradox is primarily defined by no age-differences or even superior performance on naturalistic, time-based PM tasks in older adults, but worse performance on time-based PM tasks in the lab in older adults relative to young adults (Henry et al., 2004) and old-old worse than young-old in lab, but no age differences for young-old vs old-old in naturalistic studies (Kliegel, Rendell, et al., 2008). Because time-based tasks completed in the lab offer low environmental support and require self-initiated time-checking of a clock placed out of direct sight, they are presumed to be more cognitively demanding than laboratory, event-based tasks, and consequently more sensitive to age-effects (Einstein et al., 1995; McDaniel & Einstein, 2007).

However, while most studies have demonstrated negative age-effects for laboratory time-based tasks (Einstein et al., 1995; Kvavilashvili et al., 2013; Park et al., 1997; Rose, Rendell, McDaniel, Aberle, & Kliegel, 2010), others reveal similar time-based performance for younger and older adults (d'Ydewalle et al., 1999; Logie et al., 2004; Niedźwieńska & Barzykowski, 2012). It seems likely that these discrepancies at least partially reflect the influence of artefactual variance (such as sampling error), particularly since many studies were underpowered to detect small to moderate age deficits. However, individual difference variables such as working memory capacity, or the amount of to-be-remembered information that can be maintained in the focus of attention (Rose et al. 2010), and whether the older adult sample contains young-old and/or old-old individuals (Rendell & Craik, 2000), might also potentially moderate age-effects. In addition, differences in task characteristics, such as the type of time-based task employed (Rendell & Craik, 2000; Schnitspahn et al., in press), position or visibility of the clock (Aberle et al., 2010; Niedźwieńska & Barzykowski, 2012), timing constraints of the tasks (d'Ydewalle et al., 1999), demands of the ongoing tasks (Henry et al., 2004; Park et al., 1997; Rose et al., 2010), and the demands of monitoring and judging time intervals in typical lab PM studies compared to monitoring and responding at time-of-day in typical naturalistic PM studies (Kliegel, Rendell, et al., 2008; Randall, 2016) should also be considered. For studies that have identified age deficits in lab based time-interval tasks, time management differences, which include time estimation, time-checking, planning, and plan execution, between younger and older adults appear relevant (Waldum, Dufault, & McDaniel, 2016).

Unlike a controlled lab setting, individuals can use clocks and other reminders to retrieve time-based intentions in naturalistic-settings, which likely make these PM intentions less cognitively demanding to maintain. Also time-of-day tasks in daily life have social and environmental cues indicating time-of-day—everything from position of sun in sky to what is on television and regular meal times. Indeed, one of the most robust findings in the PM literature is that most naturalistic (time-of-day) time-based PM tasks reveal positive age effects, with older adults outperforming younger adults (Henry et al., 2004; Moscovitch, 1982; Niedźwieńska &

Barzykowski, 2012; Patton & Meit, 1993; Rendell & Craik, 2000; Rendell & Thomson, 1999; Schnitzspahn et al., 2013); except when time-interval tasks are used, which appears to lead to equivalently low performance for young and young-old adults and even lower performance in old-old adults when deprived of external aids (Kliegel, Rendell, et al., 2008; Randall, 2016). Thus, the identification of positive age-effects on naturalistic time-based PM performance may reflect the historical under-use of time-interval tasks in naturalistic-settings or the over use of repeated time-of-day tasks. Specifically, many previous studies showing positive age-related effects involved time-based tasks that required participants to complete the intention multiple times throughout the observation period. There is some evidence that when individuals are required to perform naturalistic time-based tasks only once, after a delay of several days, age benefits in PM disappear (Kvavilashvili & Fisher, 2007).

Age effects in naturalistic time-based PM performance may also be influenced by the planning and specificity of the PM cues. Older adults, when planning for everyday tasks, often use time-based cues to assist in retrieval of the intention. These cues range in specificity, from broad (e.g. beginning of the week) to precise and detailed (e.g. Tuesday at 10 am). Highly specific intentions that are rich in contextual details lead to better PM performance, which may explain age-effects present in naturalistic time-based tasks (Niedźwieńska, Janik, & Jarczyńska, 2013). Furthermore, older adults may be more likely than younger adults to use contextual details when planning for future intentions and to specify more precisely when the intentions should be fulfilled, which appears to be beneficial for PM (Niedźwieńska et al., 2013). Additionally, the superior performance by older adults compared to younger adults in naturalistic-settings may also be related to the finding by Gardner and Ascoli (2015), using an experience-sampling approach, that older adults report thinking about PM tasks twice as often as autobiographical memories, compared to younger adults who reported having engaged in each memory type equally.

In summary, the PM paradox has proven to be less straightforward than originally thought, exercising the ingenuity of researchers in the PM and aging field at least since the beginning of the 21st century. While the paradox has not been fully resolved, clear progress has been made with the increasing number of PM studies using conceptually comparable tasks across settings. In particular, more event-based tasks and the two types of time-based task—time-interval and time-of-day—are beginning to appear in naturalistic-setting studies. Further investigation of the PM paradox, with both young vs older adults and young-old vs old-old adults, to disentangle the factors that contribute to inconsistent age-effects across settings will, most certainly, continue to be a fruitful area of exploration within the field of PM and ageing.

Individual Differences as Moderators of age-effects on PM

While the bulk of research investigating age-effects in PM has focused on external factors, such as the type of PM task used, there is a growing body of research that focuses on individual difference factors. In particular, changes in cognitive ability and motivation throughout the aging process—which interact in a dynamic fashion with external factors in the environment (Shelton & Scullin, 2017)—have been a topic of investigation and will be discussed in the following section.

Motivation

Motivation has been identified as an important moderating variable of age-related effects on PM performance. Aberle et al. (2010) investigated task setting and motivation as influences on PM performance for younger and older adults using a between-subject design. When motivation was manipulated using a monetary incentive younger but not older adults PM performance improved for a regular time-of-day task (sending a text message at 11am and 9pm) in a naturalistic-setting. This result may explain why young adults perform relatively poorly compared to older adults in naturalistic studies but does not shed light on two critical aspects of the PM paradox (See Age-PM Paradox section above). Firstly, the Aberle et al. (2010) findings do not explain why young adults with a strong incentive still only perform as well as older adults (rather than showing superior performance) in a naturalistic-setting; this is in contrast to the laboratory where young adults typically perform much better than older adults. Secondly, a key feature of the PM paradox not examined or accounted for by Aberle et al. (2010) is the typically substantial age-related declines from young-old to old-old on laboratory PM tasks (often larger than difference between young-old and young adults) which is typically not replicated in a naturalistic-setting (see Kliegel, Rendell, Altgassen, 2008).

In a one week diary study on the interaction of age and motivation (operationalized as perceptions of PM task importance), Ihle, Schnitzspahn, Rendell, Luong, and Kliegel (2012) found no age-related effects for PM tasks classified as “very important”.[6] However, indicative of greater motivation in general, older adults outperformed younger adults on both “important” and “less important” PM tasks. Ihle et al. (2012) found that older adults’ non-completed tasks tended to be due to self-reported “reprioritizing” while younger adults tended to be due to “forgetting”. When reprioritization was entered in a hierarchical analysis the age-related effect on PM performance was no longer significant. Although reprioritization accounted for virtually all of

[6]Issue raised post publication: *it was queried by one examiner whether this study actually implicates motivation given that there was no independent direct measure of this construct*. It should be acknowledged that motivation was only indirectly implicated or suggested by level of perceived task importance rather than being directly assessed with an independent measure. However, consistent with our interpretation of the meaning of the results, Ihle, Schnitzspahn, Rendell, Luong, and Kliegel (2012, p. 84, italics added) comment that their “results suggest that the age-related benefit observed in experimenter-given tasks transfers to everyday prospective memory and varies in dependence of *motivational* and cognitive factors.” Future research should include more independent measures of motivation.

the age-related variance it should be noted that this analysis does not control for motivation—instead it suggests that older adults are more motivated than their younger counterparts to plan their PM tasks more carefully. One potential account for this finding is that it reflects prosocial motivation (doing the experimenter “a favor”), which does appear to account for improved PM performance by older adults in the laboratory (Altgassen, Kliegel, Brandimonte, & Filippello, 2010). Taken together, it is clear that type and level of motivation are important factors to consider and investigate further in PM and aging research.

Retrospective Memory

As noted in the introduction PM is a complex construct that subsumes the retrospective memory components of encoding, storage, and retrieval (Kliegel, McDaniel, & Einstein, 2000). Empirical findings consistently show that age-related decrements in retrospective memory negatively impact PM performance. For example, Cherry et al. (2001) showed that retrospective memory accounted for 68% of age-related variance in PM performance, suggesting prospective and retrospective memory rely on similar cognitive processes that may show age-related effects (Craik & Kerr, 1996). Furthermore, increased retrospective memory load is associated with greater negative age-effects in PM (Einstein et al., 1992), with the observed decrement potentially varying with the complexity of the PM task (Kliegel, McDaniel, & Einstein, 2000)—showing the challenge of isolating the unique influence of retrospective memory on PM. Although challenging it is both theoretically and practically important for researchers to continue to investigate how dimensions of retrospective memory moderate PM performance in older adults.

Processing Speed

An influential theory of cognitive aging proposes all cognitive processes slow with aging (Salthouse, 1996). Processing speed is therefore a candidate moderator of age-related effects on PM. According to the processing speed theory (Salthouse, 1996), faster processing speed results in efficient processing of initial tasks, allowing more processing resources and time to be spent on later tasks, and better overall performance. This may explain the presence of age-related changes in more complex cognitive functions, such as PM (Schaie, 1996). During time-limited or resource-demanding ongoing tasks, PM performance may suffer, as less time encoding the intentions and monitoring for the target event is devoted to the PM task. Additionally, slower processing speed may cause the quality of information regarding PM tasks to weaken over time, making intentions less accessible (Salthouse, 1996). Supporting evidence has been provided for the relationship between processing speed and PM performance (Groot, Wilson, Evans, & Watson, 2002), as well as for the mediating effect of processing speed to explain age-effects in PM performance in older adults (Maylor, 1993; Salthouse et al., 2004; Schnitzspahn et al., 2013; West & Craik, 2001). However, there is also evidence to suggest that observed age deficits in PM cannot be fully explained by processing speed (Salthouse et al., 2004; Schnitzspahn et al., 2013; Zeintl et al., 2007).

Executive Function

Executive functioning—a complex construct involving processes such as inhibition and task-switching—is a key variable for understanding age-effects on PM. Azzopardi, Juhel, and Auffray (2015) investigated individual differences in executive flexibility as a potential mediator of PM performance in lab and naturalistic-settings. Structural equation modelling revealed age-related effects on a general PM factor—derived from both laboratory and one daily life PM measure (a call back task)—were moderated by executive flexibility, and with advancing age the influence of reduced executive flexibility on PM became more pronounced. Martin, Kliegel, and McDaniel (2003) also found that executive functions accounted for variance only in complex PM paradigms. However, this relationship was partial, as age also accounted for significant unique variance in the most complex PM tasks. Analyzing performance on a complex PM task, Kliegel et al. (2002) found 50% of age-related variance was accounted for by executive control operations. There is, however, a need for research to assess the relationship between executive functioning and PM performance in everyday life, as there have been cases of a disconnect between results of executive tests in the lab and executive functioning in naturalistic-settings (McAlister & Schmitter-Edgecombe, 2013).

Working Memory

Working memory is a limited-capacity system used to process and maintain information for short periods of time, and in PM, is theorized to be involved in planning intention execution, monitoring for PM cues, and maintaining the intention during nonrelated ongoing activities (McDaniel & Einstein, 2007; Zeintl et al., 2007). Working memory is also used for ongoing tasks, with more working memory capacity required for planning and maintaining more complex tasks. According to the dynamic multiprocess framework, individuals can flexibly switch between top-down, controlled and bottom-up, reflexive modes of fulfilling their intentions, depending on the availability of external cues, fatigue, and individual differences in processing resources (Shelton & Scullin, 2017). If environmental support in the form of external cues is not present or the association between the cue and the intention are weak, working memory resources will be needed to maintain the intention in memory.

Working memory capacity is limited, so as ongoing tasks and PM intentions become more cognitively taxing, PM performance may be disrupted (Brewer, Knight, Marsh, & Unsworth, 2010; McDaniel & Einstein, 2007). This effect is greater in late adulthood as a consequence of age-related declines in working memory (Park et al., 2002; Rose, Myerson, Sommers, & Hale, 2009; Salthouse, 1991). Several studies have shown performance on tests of working memory accounts for a significant portion of age-related variance in PM performance (Cherry & Lecompte, 1999; Kidder et al., 1997; Park et al., 1997; Salthouse, 1991). Moreover, as discussed earlier, the relationship between working memory and PM appears to be dependent on the characteristics of the PM task, such as cue focality (Brewer et al., 2010).

Clinical Applications and Issues

Research on PM and pathological aging is of critical importance for maintaining functionality in clinical groups. While healthy aging research has dominated PM research—particularly with some early work suggesting PM might be spared with ageing (see Einstein & McDaniel, 1990)—there is now a growing body of research showing robust PM deficits in a diverse range of clinical groups (e.g., schizophrenia, see Henry, Rendell, Kliegel, & Altgassen, 2007; substance abusers, see Rendell & Henry, 2009; multiple sclerosis, see Rendell, Jensen, & Henry, 2007). Some of the most ground breaking research on PM with respect to pathological aging has focused on the two most common age related neurodegenerative disorders: Alzheimer's disease (AD) and Parkinson's Disorder (PD). Hence, in this section the focus is on AD and PD while acknowledging there is a need for PM research on other age-related disorders.

Alzheimer's Disease

PM deficits are a prominent feature of AD (Farina, Young, Tabet, & Rusted, 2013; Huppert et al., 2000; Maylor, Smith, Della Sala, & Logie, 2002); evident even in the early stages of the disorder (Duchek, Balota, & Cortese, 2006; McDaniel, Shelton, Breneiser, Moynan, & Balota, 2011; Niedźwieńska, Kvavilashvili, Ashaye, & Neckar, 2017; Shelton et al., 2016; Thompson, Henry, Rendell, Withall, & Brodaty, 2010). Notably, performance on measures of PM can discriminate between healthy older adults and people with very mild AD above and beyond standard psychometric tests of retrospective memory (Duchek et al., 2006; Jones, Livner, & Bäckman, 2006; McDaniel et al., 2011).

Identifying the kinds of PM tests that are most sensitive to prodromal AD has both theoretical and practical implications. From a theoretical perspective, it is important to consider which cognitive difficulties in AD contribute to PM impairment. For example, several studies have revealed a signature decline in focal PM tasks, presumably due to disruptions in spontaneous retrieval process (McDaniel et al., 2011; Niedźwieńska et al., 2017). The observed AD-related deficit in PM tasks supported by spontaneous retrieval is particularly notable given that healthy older adults tend to perform quite well on these tasks. From a clinical perspective, it may therefore be useful to include focal PM tasks in neuropsychological assessments.

In addition to spontaneous retrieval deficits, people with AD often present with difficulty engaging top-down, controlled attentional processes (Duchek et al, 2006). Indeed, AD-related deficits have also been observed in nonfocal PM tasks in which retrieval of the intention is supported by strategic monitoring processes (Farina et al., 2013; Lee, Shelton, Scullin, & McDaniel, 2016; Tam & Schmitter-Edgecombe, 2013; Tse et al., 2015). Furthermore, performance on nonfocal PM tasks decreases as the severity of dementia increases (Tse et al., 2014), while age (e.g. young-old vs. old-old) of older adults with dementia does not play a major role in performance (Farina et al., 2013). Curiously, the ongoing cost of carrying a PM demand does not appear to differ across people with AD and healthy older adults, despite the presence of

AD-related deficits in PM performance (Lee et al., 2016; McDaniel et al., 2011; Farina et al., 2013). Thus, older adults with AD (at least in the early stages) may be monitoring for their intention execution opportunity in a similar manner as healthy older adults—but not implementing the intention as effectively. It is also possible traditional measures of strategic monitoring (i.e., ongoing task cost) could be tapping into other, unspecified, processes, thus misrepresenting potential AD-related changes in monitoring behavior. Future research could explore alternative, more direct measures of strategic monitoring (e.g., eye-tracking methods; Shelton & Christopher, 2016).

Parkinson's Disease

PM deficits have also been found with individuals with PD compared to control participants (Foster, Rose, McDaniel, & Rendell, 2013). Foster et al. (2013) suggest the onset of PD is associated with reduced executive functioning and cognitive resources necessary for effective encoding and retrieval of delayed intentions. This has been demonstrated for event-based tasks in a comparison of healthy older adults with non-demented older adults with PD, using the Virtual Week paradigm. Foster et al. (2013) found that deficits were apparent on non-focal PM tasks and on focal tasks involving greater retrospective memory demands (i.e., irregular tasks with focal target). However, a recent systematic review of PD and PM suggested that these deficits may actually be a function of co-morbidity with mild cognitive impairment (Costa, Caltagirone, & Carlesimo, 2017), the probability of which increases with age (Pankratz et al., 2015).

While at least some form of PM deficits appear to be pervasive in both PD and AD, behavioral strategies show promise in minimizing the PM-related dysfunction experienced by

PD (Foster, McDaniel, & Rendell, 2017) and AD patients (Lee et al., 2016; Shelton et al., 2016). The potential value of such strategies in the broader older adult population is discussed next.

Improving PM

External Aids and Strategy Use

A commonly-held assumption is that older adults use external aides to facilitate PM in daily life. Despite early findings that do not support such an interpretation (Rendell & Thomson, 1999; though see Experiment 3 showing older adults were more likely to keep a PDA with an alarm for time-based PM tasks closer to them than their younger counterparts), more recently this assumption has received qualified support (Masumoto, Nishimura, Tabuchi, & Fujita, 2011). External aids that assist in planning for complex PM tasks have been shown to improve older (but not younger) adults' PM performance in the laboratory (Kliegel, MacKinlay, & Jäger, 2008), but may actually be under-used by older adults in everyday PM tasks like passing on a message (Schryer & Ross, 2013). The importance of external aids and strategy use might be particularly important for older adults who have to carry out habitual (e.g., taking medication) or repeated (e.g., locking filing cabinet) PM tasks as part of their health care regime or employment duties. This is due to the robust finding that older adults are more likely to make source monitoring errors, i.e., to assume they have carried out a task when actually they have only thought about carrying out the task (Hashtroudi, Johnson, & Chrosniak, 1990; Johnson, Hashtroudi, & Lindsay, 1993; Koriat, Ben-Zur, & Sheffer, 1988).[7]

The use of external aids has typically been viewed as a confound in research on age-related effects and PM, but is an area of interest in its own right (Risko & Gilbert, 2016). Furthermore, when given the option to use an external reminder, older adults are more likely than younger

[7] Issue raised post publication: *one examiner also suggested adding the following relevant references*: Marsh and Hicks (2002) and Marsh, Hicks, Cook, and Mayhorn (2007).

adults to take advantage of the opportunity (Gilbert, 2015). Indeed, evidence suggests that individuals' confidence in their own memory abilities, as well as perceived task demands, predict the use of external aids, which, ultimately, bolsters intention execution (Gilbert, 2015). Thus, a particularly valuable direction for future research would be to pursue how older adults' confidence in, and knowledge of, their own memory (i.e., comparing perceived with actual memory performance) interacts with use of external aids.

With respect to specific types of PM strategy, the most widely studied is *implementation intentions*, which involves restating an intention in a specific format that is repeated aloud, that is a "when-then" statement: when … *cue*, then I will … *action*. (Often there is an additional instruction to imagine the situation of the intention.) For example, "*When* checking out of the hotel, *then* I will return the room key card." This strengthens the association between the cue and action, and by stating the cue first, gives priority to the cue. A large body of research has documented the beneficial effects of this strategy for a multitude of PM tasks completed in the lab and everyday life (see Chen et al., 2015, for a meta-analysis). Interestingly, the effect size associated with the strategy benefit is larger in healthy older relative to younger adults, and the enhancement observed in PM is amplified by augmenting implementation intentions with an imagination component. Not surprisingly, implementation intentions has figured prominently as a form of cognitive training to improve PM with older adults.

Training

A key distinction in the PM training literature is that between *compensatory* and *restorative training* (Reichman, Fiocco, & Rose, 2010). In practice, compensatory training typically refers to learning specific PM strategies (e.g., implementation intentions), while restorative training refers to activities theorized to stimulate neurogenesis in the neural systems supporting PM (Waldum et al., 2016). A second important distinction—related to follow up measurement of PM changes—is between the concept of *near transfer* (i.e., improvement on the same or similar tasks to those used in training at follow up) and *far transfer* (i.e., improvement on diverse PM tasks not used in training) (Rose et al., 2015). Unlike near transfer, far transfer implies generalized improvement to PM tasks of daily life—the ultimate aim of PM training with older adults. Until recently few systematic studies have been conducted on PM training with older adults, especially in terms of their long term effects (Kliegel, McDaniel, et al., 2008).

In one of the few studies to conduct such an assessment, Schmidt, Berg, Betto, and Deelman (2001) investigated whether strategy training improved PM performance and if any improvements were sustained three months later, in a middle-aged to older adult sample. Training involved using an agenda (external aid), imagery, and converting time-based PM tasks into event-based PM tasks by creating *conjunction cues*. Those in the training group performed significantly better than the control groups at the immediate follow up, but showed no difference at the three month follow up. In a more general cognitive training study with young-old adults (55-75 years of age)—including practice in attention control, task-switching, and education about matching PM strategies to particular PM task types—that also included aerobic exercise as a between-subjects variable, McDaniel et al. (2014) found an improvement from pre- to post-test[8] for the cognitive training group on VW regular and irregular PM tasks—though no difference with other groups on the time-interval (non-focal) task—compared to a control condition, aerobic exercise alone condition, and a combined aerobic and cognitive training condition. There were no differences between groups for the two other outcome measures—a health and source memory task developed for this particular study; and a simulated cooking task,

[8]Issue raised post publication: *one examiner requested the time delay between pre- and post-test for this study.* Post-test was 6 months after pre-test.

the latter developed by Craik and Bialystok (2006). Thus, benefits of training from this study are mixed, it should also be noted that there were no age group contrasts in this study, nor was age used as a covariate in analyses. Similarly inconclusive, Waldum et al. (2016), using a comprehensive and specifically focused PM training program— including education about the different types of PM tasks, their related cognitive demands, practice on a range of laboratory PM tasks, and strategies to improve PM—with a middle aged to older adult sample found an improvement in time-based PM tasks (i.e., near transfer) after controlling for baseline performance.[9] This appeared to be a function of the utilization of a learned clock checking strategy. However, for event-based tasks no difference was found between older adults in the control and training conditions.

Brom and Kliegel (2014) investigated both process (improving executive functions of task-switching) and strategy training (implementation intentions which minimized executive function demands of PM tasks) in older adults, using a naturalistic outcome measure of lapses in recording blood pressure at regular times. They found no improvement for older adults in the process training condition but a large effect for those in the training condition, suggesting that compensatory strategies may be more efficacious than process training for older adults on regular medical adherence tasks.

In sum, although valuable studies have been carried out on cognitive training to improve PM in older adults, the evidence is limited for far transfer of training gains. Thus, there is further work to be done in the area of cognitive training to improve the PM of older adults in diverse contexts.

---

[9] Issue raised post publication: *one examiner requested the time delay between pre- and post-test for this study*. There was approximately an 8 week delay from the baseline measure to the outcome measure.

Conclusions

Early empirical studies focused on PM began with an interest in how age influences this aspect of memory, and, consequently, affects autonomy and quality of life. Accordingly, early PM studies used daily life paradigms (e.g., sending a post card or calling an experimenter at a set time). With increased focus on the need to control features of the environment, these were later largely replaced by laboratory studies (in particular the Einstein and McDaniel paradigm which included a controlled ongoing task and both time and event-based PM tasks). Contrary to common assumptions, most of this literature shows that older adults typically perform well or better than younger adults on at least some PM tasks in daily life (time-of-day time-based PM tasks), and similarly on some laboratory PM tasks (event-based focal PM tasks). However, older adults do exhibit age-related difficulties on other types of PM tasks (i.e., event-based non-focal tasks, time-check tasks, and complex PM tasks). The value of ongoing PM research is identifying such vulnerabilities and finding ways to compensate for these deficits with cognitive training, and the use of behavioral strategies, and to maximize environmental support. One exciting direction is investigating PM and *future thinking* or mental time travel—to imagine and pre-experience future events. While these fields seem closely related, it is only very recently the relationship between PM and future thinking has begun to be investigated—one exception is Terrett et al. (2016) who found a significant relationship in both young and older adults.[10] There is clearly more work to be done in the exciting field of PM research with older adults, and a main aim of this article has been to spur on future research in this important area of study.

[10] To illustrate how the traditionally separate research fields of PM and future thinking are beginning to develop more explicit ties that will benefit PM and aging research, at the fifth International PM conference (ICPM) held in 2018 for the first time there were two key note presentations on future thinking. There was also a dedicated conference debate session that discussed and then voted in support for the proposal to continue the substantial inclusion of future thinking at subsequent PM conferences.

Chapter 3: Overall Thesis Methodology

In this chapter, the methods used in chapters 4 (Prospective Memory: Using parallel tasks to understand contrasting patterns of age differences seen in laboratory versus naturalistic-settings) and 5 (The role of executive control in understanding age-effects on diverse high and low demand prospective memory tasks) are presented in detail, along with relevant contextual information for each measure. The main sections of this chapter are Participants, Materials, and Procedure. The Materials section is further subdivided into PM measures, Executive Functioning measures, Other Cognitive measures, and Screening measures.

3.1 Participants

Participants in the studies reported in Chapter 4 (Experiment 2[11]) and Chapter 5 were healthy older volunteers (*n* = 104) enlisted via Villa Maria Catholic Homes' elderly independent living units, local churches, community newsletters, and word-of-mouth, as part of a larger, cognitive training parent study. The primary data used in this research project comes from the baseline phase of that parent study (Acting with The Future in Mind). Participants were compensated with $50 shopping vouchers for coming into the Australian Catholic University for all testing sessions of the larger study. Approval by the Human Research Ethics Committee of the Australian Catholic University (2015-259H) was obtained prior to commencement of testing. Participants received an information letter about the study and gave written voluntary consent at the first meeting with the researchers. Hardcopy data were de-identified (consent form stored separately to testing material) and all test forms were protected in a locked filing cabinet.

[11] Participants from Experiment 1 were from a previous research study, carried out by the second author of the manuscript submitted for publication.

3.1.1 Inclusion and Exclusion criteria

To be eligible to be included in the study, potential participants had to be over the age of 60, be in reasonable physical and psychological health (i.e., no current diagnoses of physical or mental illness), and have English as a first language. Participants were excluded if they had any history of head injuries or a psychiatric diagnosis. Furthermore, all older adults were screened for mild cognitive impairment using the Telephone Interview Cognitive test (de Jager, Budge, & Clarke, 2003), and those who had an education adjusted cut-off score of ≤ 33 were excluded (see Materials section below).

3.2 Materials

3.2.1 PM Measures

The following PM measures were selected based on their high ecological validity. All of the following PM measures have characteristics that resemble typical PM tasks that older adults might encounter in daily life. One of these measures is an established laboratory paradigm; two others are clinical measures; and finally there is a naturalistic-setting PM paradigm that is relatively new.

Virtual Week. The first PM measure chosen for this research project was Virtual Week (Rendell & Craik, 2000). It was selected as it allowed the investigation of how age-effects may appear as a function of PM task types embedded within a simulated daily life narrative. Since its development there have been multiple versions of this paradigm used with different populations in PM research (Browning et al., 2018; Habota et al., 2015; Henry, Rendell, Rogers, Altgassen, & Kliegel, 2012; Pereira, Ellis, & Freeman, 2012; Rendell et al., 2011; Rose et al., 2015; Terrett et al., 2019; Will et al., 2009). However, the first version was primarily intended to target healthy young and older adult populations (Rendell & Craik, 2000), and investigate how these two populations, but particularly older adults, may differ on simulated PM tasks of a familiar nature (e.g., pick up dry cleaning after shopping) cued by an event or time that is more or less related to the ongoing, interactive, daily life narrative, with “decisions to make and things to do” (Rendell & Craik, 2000). The environmental support of the PM tasks are manipulated by their relationship to either an event-cue (e.g., “going shopping”), a *time-of-day* cue (e.g., “4pm”), or a *time-interval cue* (e.g., 2 minutes on a stop clock), which are progressively more cognitively demanding in that order (i.e., as environmental support is expected to decrease). These differently cued PM tasks (from which they take their name) can also differ in demand in being irregular (one time only; high demand) or regular (repeated PM tasks; low demand). Furthermore, Virtual Week (Rendell & Craik, 2000; Rendell & Henry, 2009) was designed to simulate the kind of routine activities (at least for older adults), which form the backdrop of naturalistic-setting PM tasks (e.g., phone the experimenter at 10am on Tuesday). This routine and familiar background may partly account for why older adults show superior PM performance to younger adults in naturalistic-settings (Henry et al., 2004), but vice versa in the laboratory when the “classic”, rather abstract and unfamiliar, PM laboratory paradigm is used (Einstein & McDaniel, 1990; Kliegel, Rendell, et al., 2008).

The initial version of Virtual Week (Rendell & Craik, 2000) involved one practice day and then seven virtual days, or circuits of the board, with 10 PM tasks (6 regular; 4 irregular) on each of the virtual days (total of 70 PM tasks). The regular tasks included two event-based tasks (“take medication at breakfast and dinner”), two time-of-day tasks (“use asthma inhaler at 11am and 9pm”), and two time-interval tasks (“check lung capacity at 2 minutes and 4 minutes on stop clock”). This first version of Virtual Week used a physical board (31cm x 36cm), a token that the participants moved, and cards that were picked up from a pile of cards. Figure 2 shows a schematic picture of what the original version of Virtual Week looked like.

Figure 2. The original board-game layout of Virtual Week (Rendell & Craik, 2000). Copyright 1997 by Peter G. Rendell. Reproduced with permission.

Relevant for considering the two day version of Virtual Week used in the present thesis is that in this original version time-of-day was indicated by having consecutive hours of the day *marked* on squares of the board. In recent studies and this thesis, instead of having the time of day marked on squares, there is a virtual time of day clock that is calibrated to the position of the token on the board. Thus participants in this thesis had to monitor a time clock showing virtual time of day. As noted in Rose et al. (2010) the original version makes these putative time-based tasks more akin to focal event-based tasks, in which recall of the task is triggered by an event cue, i.e., when moving the token past the square with the prominent label of the time of day. Whether the qualification of focal is apt in this respect, it certainly appears to be the case that in this version the time-of-day tasks resembled event-based tasks with a salient cue. The more recent versions of Virtual Week, including the one used in this research project, include a virtual clock presented on the screen that is calibrated to the position of the token on the board.

Participants roll a traditional die, with one to six dots marked on each side (in the computer version the image of die is displayed and clicked on to activate a random number generator), and move a token clockwise across squares marked on the board depending on the number of dots shown on the die (e.g., a die roll displaying four dots would mean the participant must move the token four squares). Every two squares traversed represent "15 minutes" of virtual time. The first square corresponds to 7am and the last square represents 10pm, thus all squares represent the typical times people are awake. Some versions modify the start or finish time (e.g., Pereira et al., 2012, use 7am to 11pm), but always with the principle of normal waking hours being represented. On the circuit of the board there are 10 evenly distributed squares marked with an "E" (for event). Each time the token lands on or passes a square marked with an E, participants are required to pick up an "event" card (see Figure 3), which has a title of a virtual time relevant or typical activity (e.g., at virtual time 9am the event title is "Breakfast").

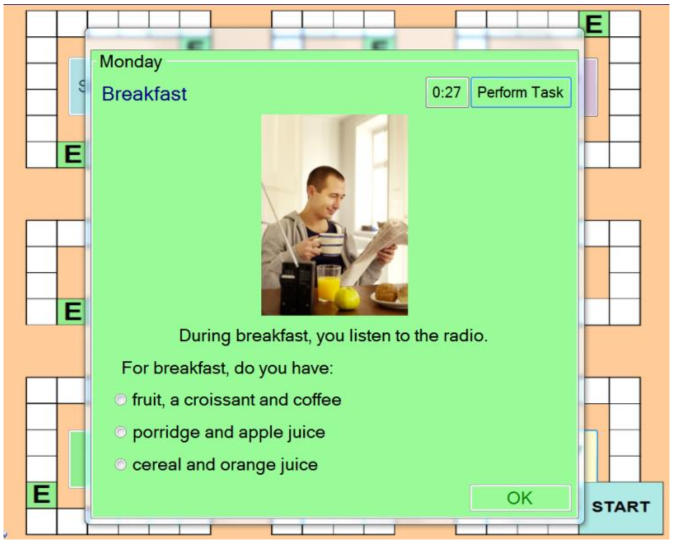

Figure 3. An example of an event card from the computerised version of Virtual Week. Copyright 1997 by Peter G. Rendell. Reproduced with permission.

The event card (or pop up screen in computer version) gives a brief description to encourage actively imagining engaging in the activity (e.g., breakfast), and presents a multiple choice question with three options of plausible and familiar activities (see Figure 3). This multiple choice decision making at event-card prompts forms part of the ongoing task, along with rolling the die and moving the token. Some of these “events” are also a cue for particular PM tasks (e.g., when picking up the event card titled “breakfast” participants are also required to indicate to the researcher that they need to perform the PM task “take antibiotics” as well as carrying out the multiple choice task of choosing what to have for breakfast). Event-based PM tasks are given either at the start of the virtual day, or contiguous with an earlier event card description (e.g., an event “you see your friend Brian”, and then a new event-based PM task, “Brian asks you to return a book to the library for him when you visit the library later that day”; later the participant will pick up an event card titled “visiting the library”).

As well as event-based tasks given at the start of the day (e.g. “take antibiotics at breakfast and dinner”) and during the day (e.g., “return book to library for Brian”) there are also “time-of-day” PM tasks given at the start of the day (e.g. “at 5pm [virtual time] phone the plumber”) or contiguous with an event card (e.g., event: “while shopping you buy a chicken”, and then new time-of-day PM task: “at 6pm put the chicken in the oven for roast dinner”). Virtual Week also includes the slightly more common time-interval PM tasks found in other laboratory paradigms (d'Ydewalle et al., 1999; Gonneaud et al., 2011; Henry et al., 2004). The time-interval PM tasks are given at the start of the Virtual Week and are regular, set times cued by a visible, and easily viewed, stop clock counting up in seconds and minutes. At the specified times participants are required to break set with the board game activity to perform the given time-interval task (e.g., “at 2 minutes and 4 minutes perform task check lung capacity”).

In the largely automated computer game versions (Browning et al., 2018; Rose et al., 2015; Terrett et al., 2019), one of which was used in the present research project, participants are required to click on a “perform task” button and select from ten options which task is to be performed. These options include various lures (i.e., plausible naturalistic tasks which were not requested to be encoded for later execution). The perform task button is to be clicked on when an appropriate event or time cue to execute a PM task occurs and before the next roll of the die in the case of event and time-of-day tasks. In the case of time-of-day tasks the cue is a virtual time of day calibrated to the position of the token on the board. Thus if a die roll takes the token past the time-of-day cue for executing the time-of-day PM task then the participant must perform the required task before the next roll of the die. In other words if, for example, a PM task is to be executed at 5pm (virtual time) and the token is currently at 4.45pm (virtual time) then, if the next roll of the die has four dots, the token necessarily has to be moved to the square corresponding to 5.15pm (virtual time). Time-interval PM tasks are classified as correct if they are performed within 10 seconds of the target time (e.g., 2 minutes on stop clock counting up). Proportion correct scores are calculated for all PM tasks and task types (i.e., number correct divided by the total possible number correct).

The two day version of Virtual Week. Particularly pertinent to the current research project are the previous studies that used the two day version of Virtual Week.[12] Ozgis, Rendell, and Henry (2009) used a two day version of Virtual Week, with two regular, and two irregular time-of-day tasks for each virtual day, with healthy and mild cognitively impaired older adults. The reason they used four rather than ten PM tasks, and two rather than seven virtual days, was due to considerations that the task might potentially be too difficult otherwise for the mild

[12] It should be noted that this refers to two *test* days, as this version, as with other versions of Virtual Week, includes one practice day where participants are gradually introduced to features of the game.

cognitively impaired group. However, unlike the two day version used in the present research project, the times were marked on squares of the board (cf. Rose et al., 2010); there was a twenty five minute filler task between the practice day and the first test day; and the irregular (i.e., to be executed only once) for both days was given at the end of the practice day. This is in contrast to normally only the regular tasks (i.e., tasks to be executed repeatedly on separate virtual days) being given at the end of the practice day. Both healthy and mild cognitively impaired older adults (in a control condition) successfully executed more of the regular than irregular PM tasks. However, another healthy and mild cognitively impaired older adult group who used the spaced retrieval technique (increasingly stretching the delay between encoding and recall of PM tasks in a rehearsal prior to the filler and test days) showed no difference between each other and on the regular versus irregular PM task dimension.

Habota, Cameron, McLennan, Ski, Thompson & Rendell (2013) used the two day computer version of Virtual Week with chronic heart failure patients and healthy controls. In this study participants used a mouse to move a cursor to click on buttons indicating components of the game (e.g., token, event-card, radio dials for multiple choice questions, etc.). The graphical user interface was designed to resemble the board game lay out used in previous tactile versions. Like the present research project, the version included a virtual time of day clock below the image of the die, with virtual time calibrated to the position of the token on the board. Participants had four regular (2 event and 2 time-of-day) and four irregular (2 event and 2 time-of-day) PM tasks. However, unlike the present research project, there was no time-interval task included and no stop watch present in the center of the board. All participants performed better on regular time compared to irregular time-cued PM tasks ($\eta_p^2 = .36$), though there was no difference in terms of the dimension of regularity for the event-based PM tasks. It appears possible that for event-based PM tasks, in the two day version of Virtual Week, the distinction between regular and irregular tasks is attenuated. In the current research project the regular and irregular dimension was collapsed and the more traditional event versus time-based task distinction (Henry et al., 2004) was focused on. Furthermore, unlike the previous studies using the two day Virtual Week version, the present research project included a time-interval (stop clock) PM task.

Reliability and validity of Virtual Week. The validity of Virtual Week has been well established across a range of populations (Rendell & Henry, 2009). The classic Einstein and McDaniel PM paradigm typically includes 3–4 PM tasks, whereas the Virtual Week paradigm includes between 4–10 PM tasks each virtual day (in the present two day version this makes a total of 20 PM tasks). Thus it is to be expected that Virtual Week will show good reliability. In a review of Virtual Week studies, Rendell and Henry (2009) reported reliability estimates ranging from .84–.94 on regular, irregular and time-interval tasks for healthy young and older adults. Using a five day, 10 PM tasks per day, computer version of Virtual Week, Rose et al 2010, found Spearman-Brown split half reliability estimates between .87–.92 for PM tasks differing on regularity and focal/non-focal dimensions.

The Memory for Intentions Test (MIST). Developed as a clinical measure of PM with high ecological validity (familiar ongoing and PM tasks), and normed on healthy and abnormally ageing older adults, the Memory for Intentions Test (MIST; Raskin, Buckheit, & Sherrod, 2010) is a widely used measure in research with clinical populations. It has an administration time of 25-30 minutes. The MIST involves participants being intermittently given a series of event and time-based PM tasks (e.g. “When I hand you a postcard, self-address it”) while they are busy working on a word search puzzle (which serves as the ongoing task). The word search form consists of a matrix of evenly spaced letters on a page. Within this matrix of letters there are 40 words (given at the bottom of the sheet) hidden horizontally, vertically and diagonally; there are also distractor words (not in the list at the bottom of the sheet). The word search sheet is also a prop for some PM tasks (e.g., writing number of medications at the top of the sheet in 15 minutes time). A digital clock is also on the table, in clear view for both participant and researcher. At the end of the PM test phase, a multiple choice recognition test of the PM tasks that were to be performed is given verbally by the researcher. The MIST also includes an optional 24 hour naturalistic-setting PM task—to call the researcher the following day at a set time and say how many hours they slept—this was not used as there was another naturalistic-setting task participants carried out as part of a new PM paradigm (see MEMO section below).

The MIST includes 8 PM trials, 4 event-based (e.g., self-addressing a postcard when presented) and 4 time-based (e.g., “In 2 minutes, ask me what time this session ends today”). Two of the event-based and two of the time-based tasks have a 2-minute delay; the other two have a 15-minute delay (e.g., “When I hand you a red pen sign your name”—the red pen is then presented 15 minutes after this instruction). Furthermore, two of each of these delays have an action response (e.g., sign name) and two have a verbal response (e.g., asking “what time does this session end?”). The scores for each trial can be either 0 (“no response”, “incorrect response to event-based cue”, or “incorrect response at incorrect time”), 1 (“correct response at incorrect time” or “incorrect response at correct time”) or 2 (“correct response at correct time [± 1 min]” or “correct response following an event-based cue [± 1 min]”). There are six subscale scores (2-minute time delay; 15-minute time delay; time-cue; event-cue; verbal response; action response) with the score of four trials being summed for each (possible range for each subscale is 0–8). The sum of the subscales gives the total raw score (possible range 0–42).

Types of errors are also recorded on a record form for administering the MIST. These are "prospective memory failure", "task substitution" (executing another intention previously encoded for MIST at an incorrect time or event cue), "loss of content" (participant indicates they "know" there is something they are supposed to do at correct time or event cue but cannot recall what), "loss of time" (participant belatedly remembers a PM task they were supposed to execute more than a minute previously), and "random error" (a completely unrelated response to any PM task previously encoded).

The researcher introduced the task using the script from the MIST manual and answered any questions the participant had. Once the participant was ready to commence, the researcher drew the participant's attention to the current time on the clock and gave the participant the first 15 minute time-cued PM task ("In 15 minutes, tell me that it is time to take a break"), the participant then immediately commenced the word search task. Using the record form for the MIST (either version A or B; these were counterbalanced due to also being used at follow up for the parent study) the researcher proceeded to give new tasks at the set times in the record form and to also present any event-cues (e.g., handing the participant a post card 15 minutes after previously telling the participant "when I hand you a postcard self-address it"). After approximately 25 minutes had elapsed (within which time all cues for PM tasks are embedded), the researcher removed the word search form, and read a series of eight multiple choice questions regarding each task cue given, to test for recognition of the correct response for each.

Figure 4 shows the series and timing of administered PM tasks given during the duration of the main part of the MIST test. The duration of the PM segment of the MIST test is given on the X axis (note that participants are informed that the test will take approximately 30 minutes, but not given any exact time for when the last PM task will be due). In Figure 4, the start of the solid bars on the left indicates when the administrator of the test needs to give the participant instructions for a PM task. The end of the solid bar on the right indicates when the previously given instructions for the PM task are to be executed (note: this may involve the presentation of an event cue as well as a time on the clock). It also shows how participants at some points during the test have to maintain multiple intentions (see where the bars overlap) while carrying out the word search task.

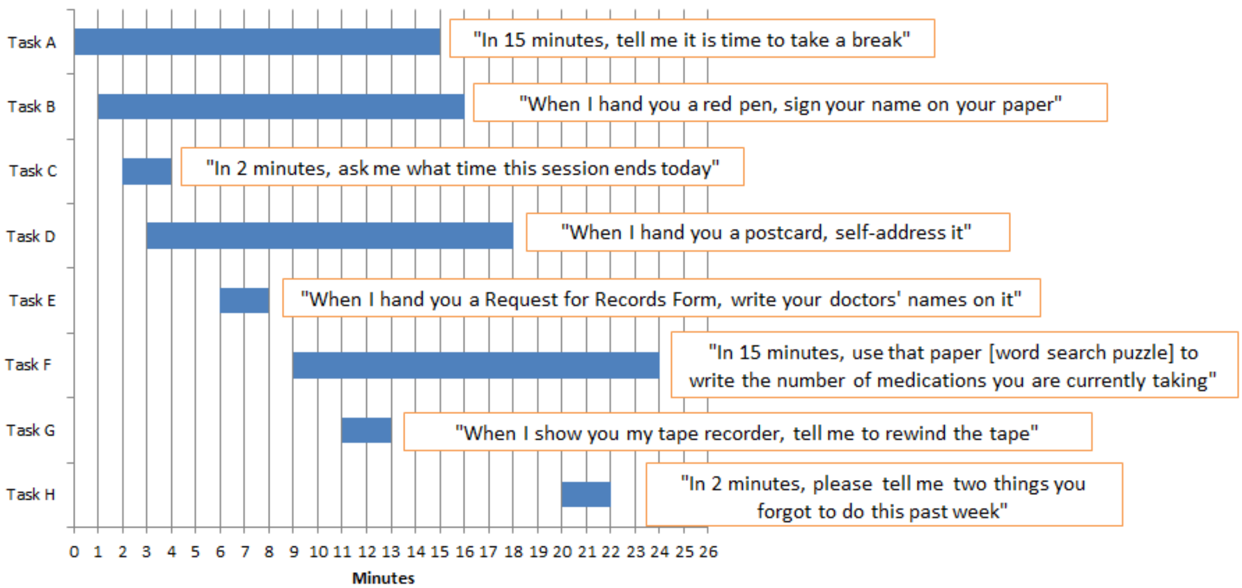

Figure 4. Duration of delay and overlap of maintained intentions for each task on MIST, with verbatim instructions for encoding of intention given adjacent (Record Form A). Event-based tasks are Task B, D, E, G; Time-based tasks are A, C, F, and H.

Reliability and Validity of the MIST. Previous research indicates that the MIST has good construct validity in terms of differential performance on event (low demand) and time-based (high demand) tasks when administered to a healthy adult sample (Kamat et al., 2014). In particular, Kamat et al. (2014) found older adults show moderate decline relative to young adults on event-based tasks (d = .41), but more robust decline for the time-based tasks (d = 1.34) which appear to rely on more self-initiated processes that decline with age (McDaniel & Einstein, 2000). Consistent with cognitive decline in self-initiated processes impacting time-based more than event-based tasks, a global executive functioning score was found to be significantly associated with time (*r* = .31) but not event-based (*r* = .20) MIST scores for older adults (Kamat et al., 2014).

When standardised with a large sample (*n* = 736) the Cronbach alpha and the Spearman-Brown split half reliability for the total MIST score were found to be high, .93 and .97, respectively (Raskin et al., 2010). The subscale alpha coefficients were .54–.64, Raskin et al. (2010, p.25) suggest that these lower alpha coefficients are "likely due to the heterogeneity of content on each trial and the limited range of scores." Intercorrelation scores for each trial using the standardisation sample ranged from *r* = .11 to *r* = .38. Interrater reliability was found to be high, with intraclass correlation coefficients ranging from .81 to 1.00, and the test-retest reliability coefficient, after a 15 day interval, was .78.

The Cambridge Prospective Memory Test (CAMPROMT). The Cambridge Prospective Memory Test (CAMPROMT; Wilson et al., 2005) is similar to the MIST in having both relatively high ecological validity and being a clinical measure of PM taking approximately 25 minutes to administer. It does, however, differ from the MIST in a number of important ways. Most notably, participants are permitted (though not explicitly encouraged) to make notes. Participants also engage in *multiple*, cognitively challenging, ongoing tasks ("puzzle sheets"), and there are *two* stop watch clocks (one for the researcher and one for the participant), as well as one analogue clock (showing hours, minutes, and seconds; set to 12pm and counting up).

Participants are instructed at the start that they will be given a number of puzzles to do (ongoing task) and that while they are doing these puzzles they "will be asked to do some other things [PM tasks] either during the session or at the end of it" (p.12). The CAMPROMT comprises six PM tasks: three event (e.g., hand an envelope when researcher says "5 minutes left”), and three time-cued PM tasks (e.g., in 7 minutes time switch puzzle sheet tasks), with equivalent delays for each type of cue. Notably, the first instruction is for the participant to observe the researcher pick up five objects and hide them in different places in the room. The participant is then asked to remind the researcher where these items are when the researcher says the words “the test is over”.

When participants fail to initiate a PM task, the researcher gives them a prompt: “you were going to do something when…” Feedback is given for each PM task response made by the participant. If the participant spontaneously performs the correct response then the feedback is simply: “good, thank you”; if an incorrect response: “yes, you had to do something but that wasn’t it, it was something else”; if the participant expresses surprise, they are given a second prompt, then if they give a correct response: “good”; otherwise: “No not quite, it was to…” and the researcher tells the participant what it was they were supposed to do. Accordingly, the scoring based on the pattern of responses on a trial for a participant are 6 (correct PM task executed spontaneously), 4 (correct response after one prompt), 2 (correct response after two prompts), and 0 (failed to carry out correct response after two prompts). As well as a total score (0-36), event-based and time-based subscales (both 0-18) are also generated. These two subscales map onto a theoretical distinction between low and high demand PM tasks which is investigated in Chapter 5 (The role of executive functioning ability in diverse PM tasks).

Reliability and validity of the CAMPROMT. The CAMPROMT event-based tasks, which are thought to involve spontaneous retrieval processes when the cue is presented, show convergent validity with the retrospective memory tasks in the Rivermead Behavioural Memory Test, r = .47; while the time-based tasks, which are thought to involve more self-initiated processes (e.g., clock checking), show less association with the Rivermead Behavioural Memory

Test, r = .22 (Wilson et al., 2005). Consistent with the idea that CAMPROMT time-based tasks require more self-initiated processing, time-based tasks correlated significantly with a measure of executive functioning from the Behavioural Assessment of the Dysexecutive Syndrome (Wilson, Alderman, Burgess, Emslie, & Evans, 1996) test, r = .50, while event-based tasks did not, r = .26. However, on another measure of executive functioning, the Trail Making Test, neither the event nor time-based tasks from the CAMPROMT were significantly correlated (Wilson et al., 2005). The CAMPROMT has shown excellent inter-rater reliability, r = .99, and reasonable test-retest reliability after a delay of 7–10 days, r = .64 when used to assess healthy adults between 16–65 years of age (Wilson et al., 2005).

MEMO. To date, the majority of naturalistic-setting PM paradigms have only included time-of-day PM tasks (see Henry et al., 2004), with a few exceptions (Bailey et al., 2010; Niedźwieńska & Barzykowski, 2012). Notably, these paradigms (e.g., making a phone call, Maylor, 1990; or mailing postcards, Meacham & Singer, 1977) and studies have been characterised by relatively few PM trials, and sometimes as low as one, which may make estimates of PM ability less stable when considering individual differences within older adult populations (for recent examples see: Bailey et al., 2010; Cavuoto, Ong, Pike, Nicholas, & Kinsella, 2015; Kvavilashvili et al., 2013; Niedźwieńska et al., 2013). Thus, there is a need for a comprehensive measure of PM task types in a naturalistic-setting with a higher number of trials. A recently developed paradigm, the MEMO (Randall, 2016), fits this requirement, with 4–8 PM tasks per day over a three day period.

The MEMO paradigm utilises a smartphone application (app) and camera function. A signature strength of the MEMO paradigm is that the event-based tasks utilise the time and date-stamp feature of a smartphone camera device. Participants receive a notification each morning on the smartphone app with a list of four previously selected, “extremely likely events” that are expected to occur that day. The event-based PM tasks are to take a photo of these events when they occur (or an honest token of their occurrence; e.g., photo of doctor’s consulting suite as an indication of the event “appointment with doctor”). Two of these events are expected to be repeated each day; with one assigned by the experimenter (take photo of lunch), and the other selected from a list of three options (brushing teeth, taking medication/vitamins or a daily landmark) by the participant. The remaining two events are unique to each day and are chosen in consultation with the participant from a list of likely events (e.g., watering garden or collecting mail). After reviewing the event-based PM tasks when opening the notification in the morning, participants press a button marked “OK” and are not shown that list again.

The MEMO event-based tasks are initially scored nominally as “correct” if the photo appears to be genuine (rather than contrived or staged). If no photo was taken of a required event-based task, and no reason given to experimenter, that, for example, it was not possible for practical reasons to execute the task, this was coded as missed. When two photos of two separate event-based tasks followed each other in quick succession (i.e., within 1 minute) the second photo was labelled as “Reminded by another task”, while the first was labelled as correct. Finally, participants were instructed that they could make a note of an event-based task which they remembered belatedly. Photos of notes to this effect were labelled as “remembered forgetfulness”. In the study reported in Chapter 5, responses were dichotomised into 1 for correct responses, and 0 for all other responses, and then totaled and divided by the number of possible correct responses to get a proportion score.

As well as for notification of the event-based (photo) PM tasks, the MEMO app also assessed two types of time-based tasks, “scheduled” and “random quizzes”. The scheduled quizzes were used to assess time-of-day PM tasks. Each morning after receiving a pop up and audible notification at approximately 8am participants nominated times for these PM tasks to occur (this is also the time [i.e., approximately 8am] when they receive the list of event-based [photo] tasks for that day). They were allowed to select from two time options for each of the tasks. The first had to be in the morning (i.e., they could choose to schedule a PM task [quiz] at either 10am or 11am) and the other in the afternoon (i.e., they could choose to schedule a PM task [quiz] at either 3.30pm or 4.30pm). As with an appointment like task, participants were required to open the app at the time-of-day they previously selected for that day (e.g., at 11am if that is the time they had selected), without any prompt from the device. These time-of-day PM tasks were scored correct if the app was opened ± 5 minutes of the required time (e.g., for a quiz scheduled for 11am the participant had to open the app between 10.55 and 11.05am to successfully complete this time-of-day PM task).

Random quizzes were the other, time-interval, time-based task measured using the MEMO. These time-interval PM tasks were designed to mimic the relatively arbitrary short delays before completing a time-based task in laboratory PM paradigms (e.g., the time-interval task in Virtual Week) and are variable in length of delay (either 10 minutes, 15 minutes, or 20 minutes). For these PM tasks, participants received a “random” notification in the morning and in the afternoon (at slightly different times each day). The notification requested the participants to open the app again after a set amount of time has elapsed. Responses are recorded as correct if they are ± 2 minute of the required time.

For both the time-of-day and time-interval PM tasks, when the participant opened the app they were required to complete a short three question “quiz”, designed to collect contextual information (obviously only for those who remember to open the app). These quiz questions had multiple choice responses and included “where are you at the moment” (with options such as at home, work, commuting, friend’s house); “what are you doing at the moment?” (e.g., relaxing, working, chores, eating), and finally a test of retrospective memory: “is this a scheduled or random quiz?”

Reliability and validity of the MEMO. There has only been one (unpublished) study that has used the MEMO (Randall, 2016), and in this initial version the event and time-based PM tasks were given on separate three day blocks (See Experiment 1 in Chapter 4). Thus there is limited data on validity and reliability. In the study by Randall (2016) correlations between the MEMO PM task types and the conceptually parallel task types in the two day version of Virtual Week, for young and older adults respectively, were .22 and .28 for event-based, .18 and -.11 for time-of-day, and .01 and .24 for time-interval. In the Randall (2016) study the Cronbach alpha for all MEMO tasks was .80 for young adults and .77 for older adults. For young and older adults, respectively, Cronbach alphas were: .71 and .66 for event-based; .68 and .57 for time-of-day PM tasks; and .57 and .74 for time-interval tasks.

In the current research project with older adults (see Chapter 5), the correlations for conceptually parallel tasks in the two day version of Virtual Week were: .14 for event-based, .07 for time-of-day, and -.05 for time-interval. Cronbach alphas were: .69 for event-based; .27 for time-of-day PM tasks; and .59 for time-interval tasks.

3.2.2 Executive Functioning Measures

The following executive functioning measures were selected due to their theoretical and empirical relationship to cognitive changes associated with ageing and PM. The constructs measured by the following tests are purported to affect older adults’ performance on PM tasks, particularly those which are high demand or require more self-initiated processes. All these tasks (and the Choice Reaction Task, see Other Cognitive Measures section) were administered on a DELL Latitude E7440 laptop computer using E-prime software (Psychology Software Tools, Pittsburgh, PA).

Task-switch task (TST). One of the facets of executive functioning ability that appears to be affected by ageing (Kray & Lindenberger, 2000) and in turn affects PM task performance (Schnitzspahn et al., 2013) is mental-set shifting, the cognitive ability to view and respond to stimuli under different aspects or rules when intermittently cued. The Task-switch task (TST) is a common and recognised measure of older adults' ability to switch between mental-sets (rules of classification). As the name implies, this task requires participants to "switch" between two cognitive tasks (Kray & Lindenberger, 2000). The TST involves three blocks of trials. In block 1 (using the version of the TST in this research project) participants are required to classify a number that appears on a computer screen as greater or less than 5 (the number 5 never appears), with one of two button presses on a computer keyboard. In block 2 participants again classify a number that appears on a computer screen but this time as odd or even, using the same two keys on the computer for the previous classification rule (the left arrow key represents "less than 5" in block 1 and "even" in block 2, while the right arrow key represents "greater than 5" and "odd", in blocks 1 and 2, respectively). In the third and final block, the rule for classifying the number appearing on the screen changes randomly during the task (i.e., after 1–7 trials, changing on average after 2 trials); with some trials requiring the participant to classify the numbers as less than or greater than 5, and some trials requiring the numbers to be classified as odd or even. The standard indices of switching ability is the difference in speed and accuracy of classifying stimuli in "pure" blocks (in which the same rule is used to classify all stimuli) compared to a "mixed" block (where the rule used for classifying the stimuli presented alternates).

Instructions were given on the screen and participants were observed during a testing phase, in which written feedback (“correct!”; “incorrect!”) was presented on the screen after each response on a trial. The researcher carefully observed the participant’s key presses on the computer to ensure that they understood which keys on the computer to press for each rule of classifying the stimuli. Particular care was taken in the third, “mixed”, block to ensure all participants understood that they needed to note the prompt presented above the stimuli (i.e. “is the number odd or even?” / “is the number greater or less than 5?”) on each trial and that they would need to go as quickly and accurately as possible. After performing the test block, participants were encouraged both by instructions on the screen, and the researcher, to respond “as fast and as accurately as you can”. Critically, in the third block participants do not know when the rule will change but are prompted as to which rule to use by a short sentence at the top of the screen (i.e., “greater or less than 5?” or “odd or even?”).

The TST provides two measures of switching costs on correctly classified trials, namely the global switch cost and the local switch cost. The global switch cost is calculated as the difference between the mean latency of the pure blocks (1 and 2) and that of the mixed block (3) and is interpreted as an index of shifting ability (the less the discrepancy between the pure and mixed blocks the better shifting ability; Braver, Reynolds, & Donaldson, 2003). The local switch cost is calculated as the difference in mean latency between repeat trials and switch trials within the mixed block. Critically, both measures of switching costs utilize only correct trial responses and do not take into account the trade-off between speed and accuracy.

The neglect of a possible trade-off between speed and accuracy is an important omission because there are often individual differences in responding, usually associated with age. That is, young adults may generally go fast on all blocks but with low accuracy in the mixed block, while older adults may go fast on the pure blocks, but slowly and with high accuracy on the mixed block. Fortunately, the TST also enables the calculation of a rate-residual score, which takes into account both accuracy and speed (Hughes, Linck, Bowles, Koeth, & Bunting, 2014). The rate-residual score is calculated using a regression equation for all participants in a given study (normative data is not currently available) and represents the difference between the observed rate of correct responding and that predicted by the regression equation. Thus, a residual that is above what is expected would indicate better shifting ability (both fast and accurate) and a residual below what is expected would reflect poorer shifting ability (fast but not accurate / or slow but accurate).

Reliability and validity of task-switch task. Global switch costs have been found to have negative correlations with reasoning ability (young adults: r = -.39; older adults r = -.23) and perceptual processing speed (young adults: r = -.61; older adults r = -.21), though only significantly so for young adults, with a similar pattern emerging for local switch costs (Kray & Lindenberger, 2000). Hughes et al. (2014) found that when errors increased on the TST, the internal reliability of the traditional latency and accuracy cost scores to have alphas of .48 and .47, respectively, while the rate residual score alpha was .72.

N-back. The n-back is a widely used measure of the facet of executive functioning ability (Smith & Jonides, 1997) referred to as “updating” (the controlled manipulation and monitoring of the contents of working memory; Miyake et al., 2000). In this version, participants are shown a series of stimuli (in this case letters), presented one at a time. For each letter, the participant must classify whether it is either the same or different to a letter that was presented a specified number of trials back (i.e., 2-back). For example, in a series beginning: A, P, W, P, … the first three letters would be classified as non-targets or different to 2-back, and the fourth would be classified as the same as 2-back target. Each letter was presented for 500 ms with a 2500 ms inter-stimulus interval. The responses in the current study involved clicking on either the right (not a target/different to a preceding letter) or left (target/same as particular letter previous in series) mouse button. An updating ability index was calculated as the average proportion of "hits" (correct responses and rejections) for targets and distractors (n = 200; 33% targets, 67% distractors). Higher proportion scores indicate better updating ability.

Participants were introduced to the concept of the 2-back (n-back) task using a diagram with different letters appearing simultaneously in a series. Practicing with the computer mouse, participants clicked on the right side of the mouse for each letter as the researcher pointed to them on the diagram, except when the letter was the same as the letter appearing two previously in the series, in which case they pressed the left side of the mouse. As there was no practice run on the computer, the researcher made sure that the participant understood the rules, including that the letters would appear only one at a time on the computer screen for a short duration, and that they would have to classify them as quickly as possible. Participants were encouraged to do their best but not to feel flustered or overwhelmed if they could not remember the letter two back in the series. In the case where they forgot, they were encouraged to guess and to renew their attention to try to pick the series up again from the letter currently on the screen.

Reliability and validity of the N-back task. The 2-back version of the n-back with older adults has been found to show convergent validity with the Trail Making Test, r = .58, though not with the putatively similar operation span tasks, digits forward and digits back span, r =-.30 and r =-.16, respectively (Miller, Price, Okun, Montijo, & Bowers, 2009). In the same study Miller et al. (2009) found the 2-back showed small to moderate correlations with the Stroop word reading, r = .17, and colour naming test, r = .43. The 2-back split half reliability coefficient has been estimated to be .85 with a young adult sample (Jaeggi, Buschkuehl, Perrig, & Meier, 2010).

Go-no go. The Go-no go is a measure of another facet of executive functioning ability, inhibition (Miyake et al., 2000), which appears to decline with age (Joly-Burra, Van der Linden, & Ghisletta, 2018). In particular, the go-no go measures the ability to inhibit a primed or pre-potent response. This speeded and repetitive task tends to generate a relatively automatic mode of responding. The task is computer-administered and consists of four blocks with 60 trials in each block. Participants are required to press the down arrow as quickly as possible whenever any letter other than X appears on the screen (the go trials). If the letter X appears on the screen (which occurs quasi-randomly on 15 trials per block), the participants are requested to inhibit their response (i.e., not press the down arrow, the no-go trials). The interstimulus interval between letters was 700 ms. Scoring involves calculating the average latency (ms) on all non-X trials and the number of errors on X or no-go trials. Inhibition ability is captured by the number of errors on X trials (the less errors the better inhibitory ability) and is calculated only for those participants whose latencies on go trials varied within a suitable range (i.e., within ± 3 standard deviations of intraindividual mean following Joly-Burra et al., 2018) to indicate they were not strategically slowing down or disengaging from the task.

Reliability and validity of the Go-no go task. The go-no go task is a widely used measure of inhibition with many variations (e.g., stimuli can be shapes or colours rather than letters, or there may be more than one target; Langenecker, Zubieta, Young, Akil, & Nielson, 2007), however, at present there does not appear to be any comprehensive investigation of the psychometric properties of a one target go-no go test (such as the one used in the present research project).

3.2.3 Other Cognitive Measures

There are other cognitive correlates of ageing besides executive functioning which are relevant for the successful performance of PM tasks. Remembering the content of a PM task is often as important as remembering that "something" needs to be done, a point captured by both of the clinical measures of PM reported above (i.e., the MIST and the CAMPROMT). It is also important to consider measures of individual differences such as intelligence when predicting performance on multiple or complex PM tasks (Cherry & Lecompte, 1999). Another relevant individual difference is perceptual processing speed (Salthouse, 1996). There is some evidence that changes in higher cognitive processes, such as PM and the putative proximal mediators of PM such as executive functioning ability, may be primarily accounted for by age related declines in perceptual processing speed. Thus, as part of this project these additional measures of cognitive functioning were collected.

Choice Reaction Task. The Choice Reaction Task (CRT) is a perceptual processing speed task. The CRT involved a presentation of blue and red squares on a computer monitor. The two coloured squares alternated intermittently, with variable inter-stimulus intervals (range 1000 to 2000 ms) to discourage anticipatory responses. The participants were asked to respond as fast as they could by clicking on the left side of the computer mouse if a blue square appeared and the right side if a red square appeared. Mean speed on correct trials was used as the index of speed of perceptual processing ability. A higher mean latency in responding corresponds to a lower speed of perceptual processing.

Reliability and validity of the CRT. The CRT has been a staple of experimental cognitive psychology research since the time of Franciscus Donders' reaction experiments in the mid-nineteenth century (Roeckelein & Roeckelein, 2006). Recent studies have found the internal reliability to be excellent, α = .94 (Deary, Liewald, & Nissan, 2011). The evidence for age related declines in processing speed on other measures with a similar format (e.g., Rose et al., 2010; Salthouse, 2000) argues for the inclusion of this measure in the current research project.

Hopkins Verbal Learning Task Revised (HVLT-R). It is often noted (e.g., Raskin et al., 2010; Wilson et al., 2005) that PM tasks can be decomposed into a prospective part (i.e., remembering something needs to be done at the right time or event in the future) and a retrospective part (i.e., the content of the intention or the actual task to be executed). Furthermore, in some cases the content of PM tasks needs to be learned (encoded) and by definition maintained during a variable interval of time. It is therefore of interest to consider how individual differences in episodic retrospective memory, particularly encoding ability and retention, correlate with PM task performance. The Hopkins Verbal Learning Task Revised (HVLT-R; Brandt & Benedict, 2001) involves encoding and recalling a list of 12 words (initial version had 15 words; Brandt, 1991), which are drawn from three different semantic categories (4 words in each category). The task is administered in four trials and involves the researcher reading a list of words to a participant. The first three trials involve immediate recall (i.e., the participants are asked to recall as many of the words, in any order, as they can immediately after the researcher finishes reading the list). The same list (in the same order) is read on each of the three trials, with a 2 second delay between each item (word) read out. The fourth, delayed and previously unannounced trial, involves the participants being asked to recall as many words as they can (from the list previously read to them in the first three trials) after approximately 20-25 minutes has elapsed. Finally, immediately after the fourth trial, a list of words including both target words (i.e., the 12 words read out in the first three trials) and semantically related and unrelated distractors, are read to the participant in a recognition memory test with a yes-no response format. HVLT-R produces four scores measuring different aspects of learning and retrospective memory: Total Recall (the highest number of words recalled from any of the first three trials), Delayed Recall (number of words recalled after delay), Retention (Delayed Recall score divided by Total Recall score multiplied by 100), and Recognition Discrimination Index (subtracting false positive responses from true positive responses). In the current research project only the Total Recall and Retention scores were used (see Chapter 5).

Reliability and validity of HVLT-R. Criterion validity has been indicated by measures from the Weschler Memory Scale-Revised, comparable to the Total Recall and Retention measures in the HVLT-R, yielding correlations of .75 and .65, respectively (Brandt & Benedict, 2001). Test-retest reliability estimates for an older adult sample tested after a six week interval was r = .66 and .40 for Total Recall and Retention, respectively (Benedict, Schretlen, Groninger, & Brandt, 1998).

The National Adult Reading Test. Developed as a measure of current and premorbid verbal intelligence (based on the observation that older adults with dementia can often read well, and that reading ability is highly correlated with intelligence), the National Adult Reading Test (NART; Nelson, 1982) consists of a series of 50 printed words that are to be read one at a time by the examinee. The words are all spelt in such a way that if they were sounded out phonetically the pronunciation would be incorrect (e.g., chord, debt, heir). The total number of errors in pronunciation can be used to look up a table in the NART manual and obtain an estimate of general and verbal IQ (the more errors, the lower the estimated IQ score; precise formula can be found in the manual).

Reliability and validity of the NART. The normative sample, 120 inpatients without brain disease, was between 18–70 years of age. Other studies have shown the NART is reliable up to 84 years of age (Nelson & Willison, 1991). Construct validity of the NART has been found with a correlation of .98 with the Full Scale IQ score on the Weschler Adult Intelligence Scale (Revised). Split half reliability and inter-rater reliability have been found to be excellent; .93 and .98, respectively (Nelson & Willison, 1991).

3.2.4 Screening Measures

Finally, to select a healthy sample of (primarily) older adults for the current research project it was necessary to include screening measures for mild cognitive decline.

The Telephone Interview of Cognitive Status modified. The Telephone Interview of Cognitive Status modified (TICS-m; de Jager et al., 2003) was administered during the participant recruitment phase to assess eligibility. TICS-m consists of 13 items measuring different domains of functioning. As the name implies, the TICS-m is administered during a phone call, to screen for mild cognitive impairment (MCI). There are four domains in the TICS-m: orientation (7 points; e.g., “what is today’s date?”); memory (e.g., immediate and delayed recall of a list of 10 words with 1 point for each of the 10 words correctly recalled for *both* immediate and delayed test; 20 points); attention (6 points; e.g. “count backwards from 20 to 1”); and language (6 points; e.g. “Please say this: ‘Methodist Episcopal’”;). Binary scoring is used (0 incorrect; 1 correct) with a maximum total score of 39 (higher scores equal better cognitive functioning). The memory domain represents 56% of the total score. The cut-off score for healthy cognitive functioning is ≥33, and participants’ raw total scores can also be adjusted for years of education (Lacruz, Emeny, Bickel, Linkohr, & Ladwig, 2013).

The Mini-Mental State Examination.[13] Designed as a measure of cognitive functioning ability, the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) consists of 11 items. Items include orientation (10 points; e.g., “what is the season?”), memory (6 points; “name 3 objects” and later same 3 item delayed recall), attention (5 points; e.g., count backwards by 7s starting from 100), with a score up to 21, and ability to name, follow instructions, and copy a complex figure, with a score up to 9. The maximum score is 30 and the cut score for MCI is 24. This measure is included here for completeness as it was used for screening participants in Experiment 1 (submitted manuscript; see Chapter 4).

3.3 Procedure

The participants were recruited as part of a larger study on cognitive training, with data from baseline testing from that study utilised in the present study. Due to the length of baseline testing as part of the parent study, two sessions were arranged for baseline assessment with each participant (approximately 1–2 weeks apart). Prior to participating in the study, participants were screened over the phone and were found to be eligible if the TICS-M was ≥33, English was their first language, and they had no history of head injuries or psychiatric diagnoses. Participants were then emailed detailed information about the parent study and arranged a time to come into the laboratory for the first session (or a time when the researcher could make a home visit if they had mobility issues). Participants were met in the foyer of the psychology laboratory located at ACU. After greeting the participant, a second information letter (the same as that previously emailed) and consent forms were given to the participants to reread and fill out once they had asked any questions and agreed to participate in the study. Participants were informed that short breaks would be available between tests, to minimise fatigue, and that the session duration would

[13] Used for screening in the collection of secondary data for older adults in Experiment 1, Chapter 4.

be between 2.5–3 hours. At the end of the first session participants were introduced to the MEMO, and given a loan smart phone with the app and camera function (further details on the induction are provided in the subsection “Session one” below). Participants were advised that they would be contacted when they returned the loan smartphone to arrange a second session which would include administration of two additional PM baseline measures (CAMPROMT and Virtual Week), and the induction into a training condition (for the parent study on cognitive training).

3.3.1 Session One

The first assessment task administered to the participants during baseline assessment was the MIST, which took approximately 30 minutes. The table was arranged (either in the laboratory or, improvising, in the participants home) so that the researcher and participant sat opposite one another. A digital clock placed at a 90 degree angle was present to keep track of time and was visible to both the researcher (either the author of the thesis; a research assistant; or one of two fourth year psychology students at ACU) and the participant. In the case of the researcher, the clock was used to time the oral presentation of new PM task instructions and visual presentation of the event cues (which were kept out of sight and removed from a participant’s view a minute after being presented). For the participant, the clock was necessary for attempts to respond correctly on time-cued PM tasks. The researcher was trained to only look at the clock discretely so as not to prompt the participant to monitor the clock or infer a PM task was due for execution. Following the MIST, a short (approximately 5-minute) break was scheduled, after which the TST was presented. The next measure administered was the NART, with pronunciation of the list of words recorded by the researcher to be scored after the session was finished. Participants were encouraged to guess if they were not sure of the pronunciation and not to feel flustered when encountering words which were unfamiliar. Following the NART, participants commenced the *n*-back. Given the cognitive demands of the *n*-back, a five minute break was called by the researcher after this test was completed.

When the testing resumed, the first three trials of the HVLT-R were administered. No mention was made of any further (delayed) testing of the words read out. The participant then commenced the TST while the researcher discretely set a timer for 20 minutes for the HVLT-R delayed recall trial. The delay of the HVLT-R was timed so that it did not overlap with any subsequent tasks, and if necessary a break was extended so that the surprise recall test could be presented in a timely manner. During the delay, participants completed the TST (approximately 10 minutes) and the CRT (approximately 10 minutes). After the delay the fourth (delayed recall) trial of the HVLT-R was administered, followed by the recognition memory test. Following this, participants completed another three tasks—grip strength measure; a continuous processing task; and timed inventory of activities of daily living—which were not used in the current research project and are therefore omitted from the Materials section. The final test to be administered during Session 1 was the Go-no go.

At the end of Session One the participants were introduced to the MEMO. The researcher gave the participant a user guide and demonstrated how to unlock the phone, and use the MEMO application and camera function. To ensure the participant understood the quiz tasks, the phone was put on test mode, and participants heard what the morning and random quiz notifications sounded like. They also practiced selecting times for scheduled quizzes and saw where the list of event-based tasks would be presented. Participants then selected events that were “extremely likely” to occur over the next three days. Once these were selected, the researcher programmed them into the phone, so that the participants would get their first set of event-based tasks the following morning, after selecting their scheduled quizzes for later in the day. Participants were also given a technical support contact number and encouraged to keep the phone with them. (In Experiment 2, see Chapter 4, participants were advised they could use “whatever strategies they wished to remember” their PM tasks.). Finally, participants were given a replied paid envelope to return the phone after the three days of MEMO administration were completed. Once the phone was received, the participant was contacted by a research assistant to make a time for the second session of baseline testing.

3.3.2 Session Two

The participant met with a different research assistant in Session Two (due to the requirements of the parent study, it was necessary for the researcher who administered the tests in Session One to be blinded to aspects of Session Two). Participants were again met either in the foyer for the laboratory or the participants’ home if they had mobility issues. The first task the participant undertook was the CAMPROMT. When the CAMPROMT was carried out in the laboratory it was relatively easy for the research assistant to plan ahead for the first PM task involving showing and hiding five objects in front of the participant (see CAMPROMT description in Materials section above). When doing a home visit the research assistant decided where to hide the object after scanning the environment. The research assistant was careful not to model note taking (which is permitted in the CAMPROMT) to the participant when discretely making a note where they had hidden the objects. After the CAMPROMT the research assistant called a short break before the next PM measure, Virtual Week, was introduced. Participants were introduced to the brief two day version of Virtual Week. Prior to the two test days, participants first did a “practice day” (circuit of the board), during which there were pop up messages gradually introducing the features of the game. Participants were requested to read all these explanatory and instructional messages out loud and to ask the research assistant if anything was not clear. Once participants finished the practice day the research assistant advised that they would only be able to give technical support and could not help the participant recall any of the required PM tasks. After Virtual Week was completed participants were inducted into a training condition for the parent study, which is not relevant to the current thesis.

Chapter 4: Empirical Paper Investigating Age-PM Paradox

4.1 Preamble

This chapter was submitted as an article for publication on 15th March, 2019 to the *Journal of Experimental Psychology: General* (see Appendix D for evidence of submission) and is presented as submitted. It comprises two experimental studies using conceptually parallel PM tasks across laboratory and naturalistic-settings. As flagged in the literature review (Chapter 2), in prior PM research there has been an asymmetry in task characteristics across laboratory and naturalistic-setting studies. This lack of conceptually parallel tasks may account for the robust pattern of age effects previously identified in this literature, known as the age-PM paradox. Indeed, as well as a lack of parallel tasks used in each setting, it has also been rare to assess PM performance of participants in both settings within the one study. Thus the primary focus of this chapter is using conceptually parallel PM task types to better understand age differences in PM. This chapter also showcases a relatively new and novel naturalistic-setting paradigm (the MEMO), which could potentially be used with diverse populations (e.g., normal and abnormally ageing older adults).

Statement of Contribution for Submitted Publication

Prospective memory: Using parallel tasks to understand contrasting patterns of age differences seen in laboratory versus naturalistic-settings.

(Submitted to Journal of Experimental Psychology: General on 15/3/2019)

I acknowledge that my contribution to the paper is 50%

Name: Simon Haines

Signature:

[redacted]

I acknowledge that my contribution to the paper is 20%

Name: Susan E. Randall

Signature:

[redacted]

I acknowledge that my contribution to the paper is 3%

Name: Gill Terrett

Signature:

[redacted]

I acknowledge that my contribution to the paper is 3%

Name: Lucy Busija

Signature:

[redacted]

I acknowledge that my contribution to the paper is 3%

Name: Gemma Tatangelo

Signature:

[redacted]

I acknowledge that my contribution to the paper is 3%

Name: Skye N. McLennan

Signature:

[redacted]

I acknowledge that my contribution to the paper is 3%

Name: Nathan S. Rose

Signature:

[redacted]

I acknowledge that my contribution to the paper is 3%

Name: Matthias Kliegel

Signature:

[redacted]

I acknowledge that my contribution to the paper is 3%

Name: Julie D. Henry

Signature:

[redacted]

I acknowledge that my contribution to the paper is 9%

Name: Peter G. Rendell

Signature:

[redacted]

Prospective memory: Using parallel tasks to understand contrasting patterns of age differences seen in laboratory versus naturalistic-settings.

(word count: 13,351)

Simon J. Haines[1]*, Susan E. Randall[1], Gill Terrett[1], Lucy Busija[2],

Gemma Tatangelo[3], Skye N. McLennan[1], Nathan S. Rose[5], Matthias Kliegel[6], Julie D. Henry[4], & Peter G. Rendell[1]

[1]Australian Catholic University, Melbourne

[2]Monash University, Melbourne

[3]Deakin University, Melbourne

[4]University of Queensland, Brisbane

[5]University of Notre Dame, South Bend

[6]University of Geneva, Geneva

*Correspondence

Simon J. Haines

simon.haines80@gmail.com

Abstract

Prospective memory (PM) affords autonomy and quality of life in older age. Prior PM research shows paradoxical findings whereby young adults outperform older adults in laboratory-settings but the reverse is found naturalistic-settings. Moreover, young-old outperform old-old adults in laboratory-settings, but show no age differences in naturalistic-settings. However, task characteristics have differed between laboratory and naturalistic-setting studies. Thus, the paradox of contrasting PM performance between age groups across settings is potentially a function of comparing disparate task types. In two experiments, we tested this hypothesis using analogous paradigms across laboratory and naturalistic-settings, with three types of PM tasks: event-based, time-of-day, and time-interval. Experiment 1 compared young (*n* = 40) and older (*n* = 53) adults on a laboratory paradigm that measured PM tasks embedded in a virtual, daily life narrative; and on a conceptually parallel paradigm using a customized smartphone application (MEMO) in actual daily life. Results revealed that on the MEMO, older adults outperformed young adults on the time-of-day tasks but did not differ on the time-interval or event-based task. In contrast, older adults performed worse than young adults in the laboratory. Experiment 2 compared PM performance in young-old (*n* = 64) and old-old (*n* = 40) adults using the same paradigms. Young-old outperformed old-old adults in the laboratory, however, group differences were not evident in daily life. These findings show the limiting conditions of the age-PM paradox, the need for a finer theoretical delineation of time-based tasks, and the potential for MEMO to clarify the interaction of levels of environmental support with PM in naturalistic-settings.

Keywords: Prospective memory; Cognitive Age-Effects; PM Paradox; Ecological Validity

Understanding the effects of normal aging on prospective memory (PM)—the ability to remember and execute delayed intentions at an appropriate time or event in the future (Ellis, 1996)—is important for providing appropriate support for independent living in older age (Haines et al., 2019; Hering, Kliegel, Rendell, Craik, & Rose, 2018; Kliegel et al., 2016; Lee, Ong, Pike, & Kinsella, 2018). This is apparent when considering common PM tasks, such as collecting medication from the pharmacy and taking a meal out of the oven. Currently, however, a clear picture of the effects of normal aging on PM has not been established, largely because a number of apparently contradictory findings have been reported regarding the PM performance of young and older adults across laboratory and naturalistic-settings; a pattern that has been referred to as the age-PM paradox (Rendell & Thomson, 1999). The first key aspect of the paradox relates to the observation that, when comparing group means, young adults typically outperform older adults in laboratory-based studies, while the reverse is observed in naturalistic studies (Henry et al., 2004; Kliegel et al., 2016). The second key aspect of the paradox, and one that is relatively neglected in the literature, relates to the pattern of PM performance of young-old and old-old adults. Specifically, young-old adults have been shown to outperform old-old adults in laboratory-based studies, but curiously age differences are absent in naturalistic-settings, with both groups performing relatively well (Kliegel, Rendell and Altgassen, 2008; Rendell & Thompson, 1999; Rendell & Craik, 2000).

A range of factors have been proposed to explain these age-effect discrepancies across laboratory and naturalistic-settings, including variations in cognitive demands of the ongoing task (i.e., whatever additional activity is being performed during the delay period and opportunity for intention execution; Schnitzspahn et al., 2011), and age-related differences in the use of external aids like dairies or alarms in naturalistic-settings (Ihle et al., 2012; Maylor, 1990;

Moscovitch, 1982). However, none of these factors as yet has been shown to adequately account for the paradoxical findings. In the present experiments we investigated another possible contributor. Specifically, we tested whether both aspects of the paradox may in fact reflect the different methodologies that have typically been used across laboratory and naturalistic-settings, and as such whether the paradox may be reduced—or eliminated—by using conceptually similar (i.e., analogous or parallel; though not literally the same) PM tasks in both settings.

When attempting to unpack the PM paradox and to document age-effects on PM, one key feature to consider is PM task type (Niedźwieńska & Barzykowski, 2012). According to the multiprocess framework of PM (McDaniel & Einstein, 2000), in general, event-based tasks (e.g., passing on a message when you see a colleague) involve relatively spontaneous cognitive processes that tend to be spared with normal aging (Einstein & McDaniel, 1990). Thus, while performance on event-based tasks may be expected to decline with age, this reduction may be relatively subtle. On the other hand, time-based tasks, particularly time-interval tasks that involve monitoring the passage of time (e.g., taking cakes out of the oven in 20 minutes) are assumed to involve more executive, effortful cognitive processes (Hasher & Zacks, 1979), which tend to be negatively affected by normal aging (McDaniel & Einstein, 2008). Thus, more substantial age-related decline in performance might be anticipated. Theoretically, however, it should be noted that one type of time-based task, namely *time-of-day* tasks (e.g., phoning the doctor at 1pm) performed in daily life, might be anticipated to show age-effects that fall somewhere between these two extremes, as they lend themselves to being associated with various environmental events and cues that might be expected to occur at a similar time. For example, 1pm is a common lunchtime, so eating lunch might act as a cue that it is 1pm, and that you need to call the doctor. These potential opportunities for environmental support are known as *conjunction cues* and may make time-of-day tasks more similar to event-based tasks than time-interval tasks (Maylor, 1990; Phillips et al., 2008; Rendell & Craik, 2000).

The distinction we are drawing here for time-based tasks in many ways parallels, and is of similar importance, to a distinction frequently made within event-based tasks (Haines et al., 2019; Kliegel, Jäger, et al., 2008). This distinction is between event-based cues which are focal (where the cue is related to the cognitive processing involved in the ongoing task) and event-based cues which are non-focal (where the cue is unrelated to the cognitive processing involved in the ongoing task). Focal event-based tasks are assumed to involve relatively spontaneous processes, less affected by aging (Einstein & McDaniel, 2005). In contrast, non-focal event-based tasks are assumed to involve relatively effortful, strategic cognitive processes, which are markedly affected by aging (McDaniel & Einstein, 2007). We propose that, within the category of time-based PM tasks, "time-of-day" tasks, which are relatively high in environmental support, are analogous in terms of cognitive demand (and hence age-effects) to focal event-based tasks, while "time-interval" tasks, which are relatively low in environmental support (assuming no external aids are used), are analogous in terms of cognitive demand (and age-effects) to non-focal time-based tasks. This distinction between time-based tasks has been highlighted previously (Haines et al., 2019; McDaniel & Einstein, 2008; Randall, 2016). For example, in a chapter on PM by McDaniel and Einstein (2008) they write: "In *time-based* prospective memory tasks, the appropriate moment for executing the prospective memory intention is a particular *time-of-day* (a doctor's appointment) or *the passage of a particular amount of time* (taking cookies out of the oven in 10 minutes)." (Emphasis added.) However, the significance of this distinction has received little research attention in the published age-PM literature (Haines et al., 2019).

In sum, according to the multiprocess framework, age-effects on PM are a function of PM task characteristics (e.g., event, time-interval, or time-of-day tasks), such that older adults may be more vulnerable to PM lapses when either the cognitive demands of a PM task increase (e.g., needing to interrupt an engaging activity intermittently to check a clock) or environmental support decreases (e.g., less cues in the environment that may act as reminders; Craik, 1986).

Given this, one potential and largely unexplored reason for the age-PM paradox pattern of results in the literature could be the use of conceptually disparate PM tasks across laboratory and naturalistic-settings (Phillips et al., 2008; Rendell & Craik, 2000; Schnitzspahn et al., 2011). In particular, naturalistic measures of PM are predominantly time-of-day tasks (Henry et al., 2004; Kvavilashvili & Fisher, 2007; Maylor, 1990; Schnitzspahn et al., 2011) designed to resemble scheduled appointment-like tasks (e.g., call experimenter at 4pm), while laboratory measures of PM are primarily event-based tasks (e.g., press a button whenever an animal word appears; Henry et al., 2004). Furthermore, when time-based tasks are used in the laboratory, they tend *not* to be time-of-day tasks, but are more typically time-interval tasks which require an action to be performed after a short time has elapsed (e.g., press a button after two minutes). In addition, the difference between time-based tasks often used in the lab and those commonly used in naturalistic-settings is accentuated by the fact that there are many features of daily life that indicate time-of-day, and can therefore serve as conjunction cues for time-of day PM tasks in naturalistic-settings (e.g. clocks, TV or radio programs, level of sunlight, foot and vehicle traffic, other people eating meals or snacks, etc.). Such features may therefore facilitate performance on this type of time-based task in naturalistic-settings (see Henry et al., 2004), and these cues may be more reliable for older adults who lead more structured or routine lives compared to the socially unpredictable lives of young adults (Carstensen, Isaacowitz, & Charles, 1999).

Relatively few studies have tested PM in laboratory and naturalistic-settings within the one study, and the use of parallel tasks in each setting is even rarer. Work by Rendell and Craik (2000) provides one of the few attempts to date to systematically address the issue of whether the key features of the age-PM paradox are still apparent when PM task types are designed to be conceptually similar across laboratory and naturalistic-settings. In their study, Rendell and Craik tested performance in the laboratory using Virtual Week, a laboratory paradigm in a board game format that simulates the typical structure and activities of daily life, and involves an ongoing activity of making various choices throughout the game (for more recent versions see: Browning et al., 2018; Rendell & Henry, 2009; Terrett et al., 2019). A range of PM tasks (including virtual time-of-day tasks) are embedded in the game. The paradigm Rendell and Craik (2000) used for the parallel tasks in a naturalistic-setting was Actual Week, which involved completing the same number and type of PM tasks per day as in Virtual Week (i.e., 10), at set times or in relation to set events within the participants' normal daily life routine.

The Rendell and Craik (2000) study largely confirmed the general paradox, finding that young adults were superior to older adults in the laboratory-setting (Virtual Week) but inferior in the naturalistic-setting (Actual Week); and young-old were superior to old-old adults in the laboratory but showed no age differences in a naturalistic-setting. There was one exception to the typical paradox pattern: young-old were superior to old-old adults on the time-interval task in Actual Week (young-old were also superior to the young adults, consistent with the first aspect of paradox). Thus, the age-related decline seen in older adults in laboratory-settings was partially replicated in the naturalistic-setting for the time-interval task but not for the time-of day or event-based tasks. However, two key limitations of the Rendell and Craik (2000) study, which the present study seeks to address, need to be noted. First, in Actual Week successful performance on the event-based tasks could not be adequately verified. More specifically, while the recording device could verify when the time-based task was completed, it was not possible to verify whether the event-based cue had actually occurred, or whether the PM task had been completed as instructed. The second key limitation of Rendell and Craik (2000) was that, using the first version of Virtual Week, the validity of the time-of-day tasks may have been compromised (cf. Rose et al., 2010). While recent computer versions of Virtual Week have a virtual time clock calibrated to the token position on the game board (e.g., Terrett et al., 2019), the original manual version of Virtual Week had the consecutive hours of the day marked on the squares on the board, which may have acted like event-based cues when the token passed the marked square corresponding to the set time. Indeed, in a Virtual Week study by Rose et al. (2010) the version of time-of-day tasks involving marked squares was classified as a type of event-based task rather than a time-based task.

In a more recent attempt to investigate whether the use of disparate PM task types drives the age-PM paradox, Bailey et al. (2010) created parallel forms of an event-based task across laboratory and naturalistic-settings. Adapting smartphone technology Bailey et al. (2010) used signal-contingent experience sampling to manipulate ongoing task characteristics (a survey of 45 questions three times a day for five days) in a naturalistic-setting, such that the task more closely resembled standard laboratory ongoing tasks. The PM cue was structural (i.e., a question in uppercase letters) and the PM response was to break set from using the 1-5 keys for responses to questions and instead press the # key. Using this paradigm in a naturalistic-setting, Bailey et al. replicated age-effect deficits that are typically found in laboratory-settings. This suggests that the highly controlled, unfamiliar and artificial aspects of the laboratory situation may differentially and negatively affect older adults' ability to plan and execute PM tasks. Converging evidence for this possibility was found in a laboratory experiment that required participants to execute errand planning tasks in an imagined naturalistic-setting (i.e., going to a shopping center and carrying out various errands), and perform theoretically similar errand planning tasks in an imagined unfamiliar, science fiction scenario (i.e., travelling in a spaceship and visiting different planets to perform various fantastical errands; Kliegel, Martin, McDaniel, & Phillips, 2007). While young adults managed to successfully schedule the majority of errands in both imagined settings, older adults' performance was comparable with young adults only in the naturalistic scenario, and was substantially reduced in the more artificial, science fiction setting. Given these findings, which highlight the apparent influence of the artificiality of laboratory tasks to date, the laboratory task selected for use in the present experiment was a recent version of Virtual Week (VW; Rendell & Craik, 2000; Rendell & Henry, 2009) that was modified to overcome the limitations in the validity of time-based tasks in the original version of Virtual Week (Rendell & Craik, 2000) outlined previously.

A recent attempt to create conceptually parallel time-based PM tasks across settings was devised by Schnitzspahn et al. (2011). In that study the time-based tasks in the lab involved pressing the “a” key on a computer keyboard at two set times within three 12-minute blocks of a naturalistic ongoing tasks (i.e., watching documentaries). The duration of the time-based tasks in the lab were scaled up in the naturalistic-setting, with one set time cue in the first 12 hours of the day, and one in the last 12 hours, on three consecutive days. The time-based task in the naturalistic-setting was to send two text messages, with “a” as the content, in the morning and afternoon. However, a limitation of this approach is that the lab time-based task appears to be a time-interval task (few or no conjunction cues possible) while the apparent parallel task in the naturalistic-setting appears to be a time-of-day task, and therefore amenable to a rich array of conjunction cues. Understandably practical limitations of the laboratory-setting may play a role in the challenge of developing conceptually parallel time-of-day tasks across settings. The Virtual Week paradigm provides an elegant solution to the difficulty of separating time-of-day and time-interval tasks by providing, in a virtual format, some of the conjunction cues (e.g., regular meal times) that occur for time-of-day tasks in naturalistic-settings, while still including the more common lab, time-interval task; involving breaking set with the ongoing task (i.e., switching attention from engaging in the virtual day with its various plausible commonplace scenarios).

The Current Experiments

The aim of the current research was to investigate the extent to which both aspects of the age-PM paradox (i.e. young vs old, and young-old vs old-old patterns of PM performance) reported in the literature are a reflection of the lack of conceptual comparability of PM tasks across laboratory and naturalistic-settings. To do this we built on, and substantially improved the previous work by Rendell and Craik (2000) by first of all addressing the limitations of Actual Week. We did this by developing a novel, naturalistic-setting PM paradigm which collected data on PM performance during participants' everyday lives via a mobile phone application (entitled MEMO). The PM task types were conceptually equivalent to those in the recent version of the Virtual Week laboratory paradigm which was also used in this study. More specifically, we incorporated into MEMO event-based and time-of-day tasks similar to the tasks administered in prior studies, as well as time-interval tasks (e.g., complete a task after 5, 10 or 15 minutes has elapsed). The event-based tasks in MEMO involved taking digital photos in response to particular event cues (e.g. at lunchtime take a photo of your meal). As such, in the current experiment, the event-based tasks were able to be verified as having been carried out in response to the set event cue by checking both the time-stamp *and* the content of the photos. In addition, administering MEMO via a mobile phone app offered a further improvement on Actual Week, which had required participants to carry around, and have access to, a cumbersome portable recording device. By contrast, a mobile phone is a familiar device and generally kept accessible by most people. This helps address the possibility that young adults performed poorly relative to older adults in naturalistic-settings—such as in Rendell and Craik (2000)—because they had more difficulty keeping the recording device close to hand. We also addressed limitations regarding the laboratory Virtual Week paradigm. Specifically, we addressed the problem of the virtual time-of-day tasks potentially having features of event-based tasks (i.e. triggered by the event cue of passing the marked square and thus reducing or eliminating the need to monitor the virtual time). To do this, we used Virtual Week with the recent innovation (e.g. Rendell et al., 2011) of a virtual time clock calibrated to the token position on the board, which replaced times marked on the board. Thus, our version of Virtual Week required participants to monitor the virtual time of day by actively checking the virtual time clock.

It should also be noted that an additional advantage associated with the use of MEMO in comparison to Actual Week was that the time-based task we chose to include in MEMO was to respond to a short survey at a predetermined time. Unlike Bailey et al. (2010), we did not embed the PM task within an experimenter-given ongoing task (i.e., responding to a long list of questions and then switching response when the PM cue appeared). Instead, completing the survey itself constituted the PM task, and the normal activities of daily life acted as ongoing tasks. Furthermore, participants' answers to the survey gave us the secondary advantage of capturing some information about the ongoing activities that individuals were engaged in within their actual daily lives when they remembered to perform the PM task. This additional information is valuable because although ongoing tasks in naturalistic-settings (e.g., visiting a shopping center) can be simulated in the laboratory with a paradigm like Virtual Week, it cannot be assumed that these simulated tasks will necessarily correspond to the actual ongoing tasks of individuals at different ages. Furthermore, the bonus contextual information afforded by the MEMO paradigm may help address the question of potential systematic differences in the naturalistic-setting contexts within which young and older adults successfully execute PM tasks. This is valuable as some commentators of the PM paradox have speculated that differences in the demands of the ongoing tasks that young (more demanding ongoing tasks) and older adults (less demanding ongoing tasks) are typically engaged in outside the laboratory may explain the PM paradox (Phillips et al., 2008).

Adopting these methodological improvements to PM assessment tasks, we set out to undertake the most comprehensive investigation to date of the extent to which the lack of conceptually similar PM tasks across laboratory and naturalistic-settings might account for the two key aspects of the age-PM paradox. To do this, we conducted two experiments. Briefly, the first experiment addressed the paradox in relation to PM performance of young compared to older adults across settings (young adults superior to older adults in laboratory-settings but the reverse pattern in naturalistic-settings), and the second experiment adopted a similar approach to address the paradox in relation to the PM performance of young-old versus old-old adults (young-old superior to old-old in laboratory-settings but no age differences in naturalistic-settings).

Experiment 1

Experiment 1 used a repeated measures design with a sample of young and older adults. It incorporated analogous types of PM tasks in naturalistic and laboratory-settings. In the naturalistic-setting MEMO was used as the PM paradigm. MEMO was specially developed for the purpose of creating conceptually parallel PM task types in naturalistic-settings to those used in a laboratory-setting by the Virtual Week paradigm. Based on a large body of previous laboratory research using Virtual Week (Henry et al., 2004; Rendell & Henry, 2009) it was hypothesized that, in the laboratory-setting, young adults would perform more successfully than older adults on all three types of PM task (i.e., event-based, time-of-day, and time-interval), but that these group differences would be greatest for the (virtual) time-of-day and interval time-based tasks. (It should be noted that although time-of-day tasks may resemble event-based tasks, which may facilitate spontaneous processes of cue recognition, that this is more so in terms of salience of (temporal) cues, than in terms of cue focality, i.e., overlap with the ongoing task. This means that in the laboratory, where there are less natural, or potential for simulated, temporal cues, older adults would be expected to show markedly lower levels of performance compared to younger adults than is expected to be the case with event-based tasks in Virtual Week). For the analogous PM tasks administered in a naturalistic-setting, the hypothesized pattern of results was that older adults would show superior performance on the event-based and time-of-day tasks (which lend themselves to environmental support) in comparison to young adults. However, for the time-interval tasks, it was hypothesized, based on the findings of Rendell and Craik (2000), who used 30 and 60 minute delays, that young and older adults would show more commensurate performance to that evidenced in the laboratory on the time-interval tasks, i.e., both groups would show relatively poor performance compared to the event-based and time-of-day tasks as they are thought to be most taxing of cognitive resources due to low environmental support and high strategic monitoring demands.

Method

Participants

Ninety-three healthy, community dwelling volunteers—40 young (19–30 years; 75% female) and 53 older (65–86 years; 68% female) adults—were recruited via flyers at Australian Catholic University, recreation facilities, churches, the University of the Third Age, and a Probus group. Participants received $30 for their participation, with the exception of undergraduate students, who obtained partial course credit. Older adults were screened using the Mini-Mental State Examination (M = 28.85; SD = 1.28), and those who scored below 24 (Folstein et al., 1975) were excluded. Twelve older adults were excluded from analyses. Of these, five declined to complete the experiment, five completed less than one day of the naturalistic phase of the experiment, and two others reported neurological conditions, leaving a total of 40 young and 41 older adults whose data could be included in analyses. Characteristics of the two age groups are reported in Table 1.

Table 1
Characteristics of Participants in Experiment 1

	Young Adults		Older Adults		*t-test*		
	M	*SD*	*M*	*SD*	*t*	*p*	*d*
Age (in years)	24.13	3.63	71.61	4.86			
Education (in years)	16.34	2.61	14.55	3.04	2.84	.01	0.63
Verbal IQ[a]	102.15	8.68	110.00	6.31	4.66	< .001	1.03

[a]Estimated using the National Adult Reading Test (NART; Nelson, 1982).

Materials and Procedure

Participants were tested individually in a 2-3 hour laboratory testing session at the Australian Catholic University[14] in Melbourne, followed by six days of naturalistic testing (three days of event-based tasks and three days of time-based tasks, counterbalanced). To minimize fatigue, breaks were provided as needed throughout the testing session in the laboratory. The laboratory testing session entailed participants completing a demographic questionnaire, followed by the Mini-mental State Exam (MMSE; Folstein et al., 1975; for older participants only) and the National Adult Reading Test (NART; Nelson, 1982). Participants were then instructed on how to use the naturalistic measure of PM (MEMO) before the laboratory PM measure (Virtual Week) was administered. Prior to leaving the laboratory, participants were again briefed on the MEMO.

Laboratory PM measure. Virtual Week (VW; Rendell & Craik, 2000; Rendell & Henry, 2009) is a comprehensive measure of PM presented in a board-game format, that has been adapted for computer automation (Browning et al., 2018; Henry, Rendell, Phillips, Dunlop, & Kliegel, 2012; Leitz, Morgan, Bisby, Rendell, & Curran, 2009; Mioni, Rendell, Stablum, Gamberini, & Bisiacchi, 2015; Niedźwieńska, Rendell, Barzykowski, & Leszczyńska, 2014; Rendell, Jensen, et al., 2007; Rendell et al., 2011; Terrett et al., 2015). Figure 5 shows the board game image, as well as examples of images that participants see when they are prompted to select task types and when they perform PM tasks. To complete Virtual Week, participants must click on the image of a die (a random number generator of 1 to 6), and move a token around the board according to roll of the die, with each two squares traversed on the board representing "15

[14] One young adult and one older adult preferred to be tested at their homes. The data were collected by the second author as part of her PhD and also involved a manipulation of encoding using implementation intentions. No systematic significant differences were found for the implementation intentions groups—possibly due to participants not using the strategy as requested. However, as implementation intentions is not the focus of the present experiments, control and implementation intentions groups for each age group were collapsed for all analyses.

minutes" of a virtual day. A virtual day is represented by a complete circuit of the board (60 squares), beginning at "7am" and finishing at "10pm". A virtual clock, which changes as the token is moved, and a stop clock, counting up by seconds and minutes in real time, are displayed in the center and at the top of the screen. They are relevant for the time-of-day and time-interval PM tasks, respectively. There are 10 event squares marked on the board. When the token passes one of these participants are required to click on the button labelled Event Card to reveal the image of an event card. The event cards form a plausible daily narrative, which acts as the ongoing task (e.g., "you visit the doctor") and participants are required to make choices at each event (e.g. "while waiting for the doctor you read: a) a magazine; b) a book you brought with you; or c) some brochures"). A button marked "Perform Task" is always present in the top right-hand corner of the board and at a similar location on event cards. Participants must click this button to perform each of the pre-determined PM tasks. The "Perform Task" button can be pressed at any time (or event) and offers a menu of 10 PM tasks that may be selected to be performed. Participants are advised that they can still perform a task even if they realize it is late or the appropriate opportunity has passed.

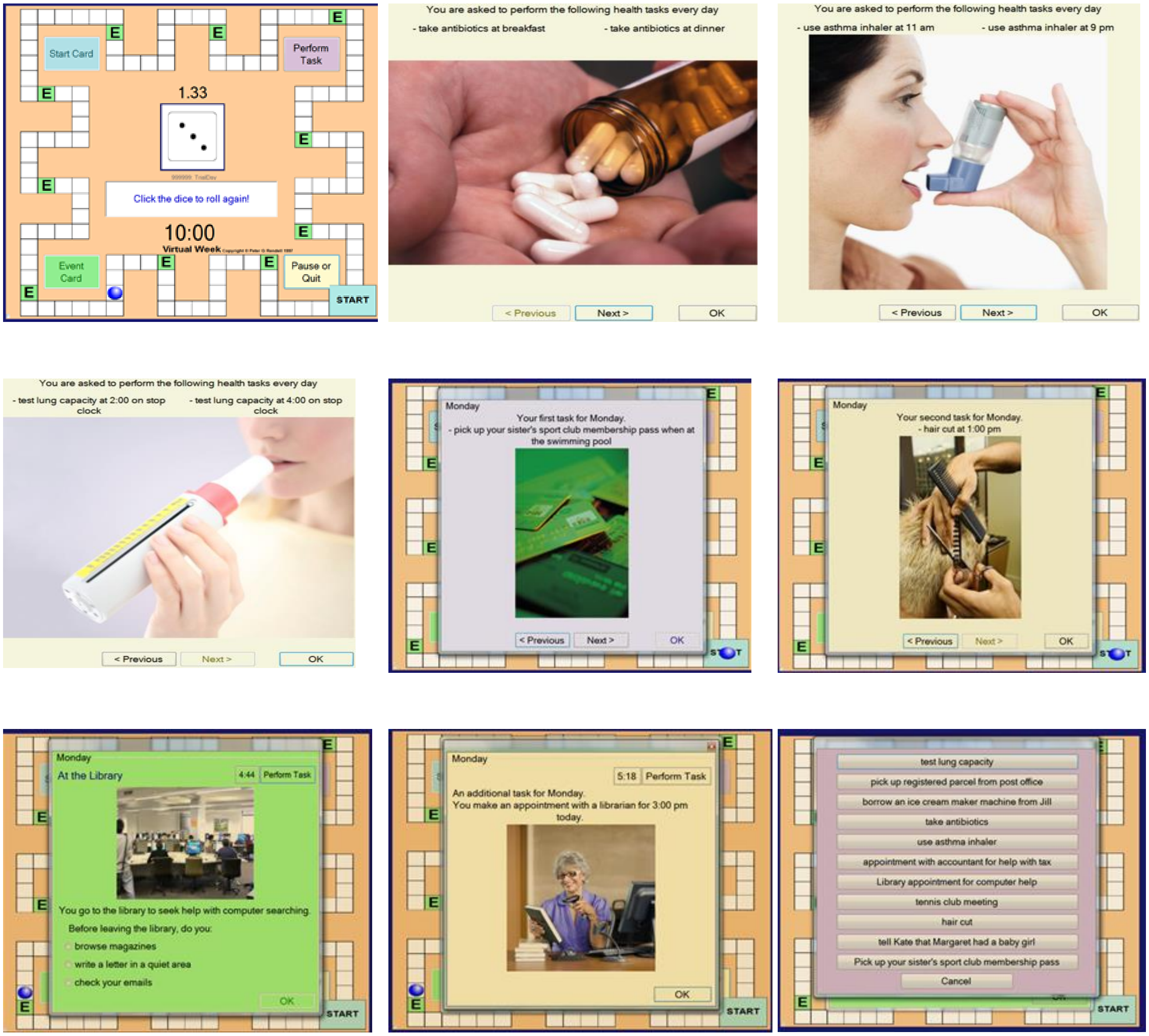

Figure 5. Screen shots from Virtual Week computerized version. From left-right, top row: board game interface; event-based task ("take antibiotics at breakfast and dinner"); time-of-day time-based task "use asthma inhaler at 11am and 9pm"); middle row: time-interval task ("test lung capacity"); irregular event task start of day ("pick up … at swimming pool"); irregular time-of-day task, start of day ("hair cut at 1pm"); bottom row: ongoing task event card ("At the library"); additional irregular time-of-day task during day ("You make an appointment … 3pm"); perform task menu. *Copyright 1997 by Peter G. Rendell. Reproduced with permission.*

At the beginning of the first day participants are given regular event- and time-based tasks, meaning the same tasks are to be performed each virtual day and always in response to the same cues. Additionally, at the start of each day, before rolling the die, participants are given two "irregular" PM tasks unique to that day (one time-of-day—e.g., "haircut at 1pm"—and one event-based—e.g., "return book when at the library"). Two additional irregular tasks (one time-of-day, one event-based) are presented to participants after particular event-cards have been picked up during the virtual day (e.g., after interacting with an event-card about dropping off photocopying, a task card pops up with the following new PM task "remember to call your partner at 4pm to collect photocopies" [time-of-day]; after being informed in an event card that their friend had a baby girl, a task card pops up with the following new PM task "when you see Margaret remember to tell her about Jane's new baby" [event-based]). As illustrated with the preceding examples, these PM tasks that "pop up" during the virtual day follow on from the events described in the event cards, adding to the realistic narrative quality of Virtual Week.

The first circuit of the board is a practice day, in which participants are introduced to features of the game, become familiar with the board game format, and can ask the experimenter any questions. The experimenter is actively involved in guiding participants through the practice day which also includes automated help messages and instructions. However, this guidance by the experimenter is limited only to technical support, and the presentation of help messages cease once the actual PM assessment task begins. Following the practice day, participants completed 2 virtual days with a total of 20 PM tasks: 8 event-based tasks (4 regular, 4 irregular); 8 time-of-day tasks (4 regular, 4 irregular), and 4 time-interval tasks (all regular[15]; i.e., “check lung capacity at 2 and 4 minutes on the stop clock”). After the practice day the experimenter remained present but the participants completed the virtual days by themselves without any further guidance by experimenter.

Scoring criteria. The proportion of correct responses for each of the PM task types (regular/irregular event-based, time-of-day, and time-interval tasks) was used as an indicator of PM performance. A correct response was defined as one where the participant completed the appropriate PM task in response to the relevant cue. More specifically, responses were categorized as correct if they were performed when the token arrived at (or just passed) the target position on the board, and before the next die roll. In regards to time-interval tasks, responses were categorized as correct if the task was completed within 10 seconds of the target time.

Naturalistic PM measure. Developed as a parallel measure for the PM task type data captured by Virtual Week, MEMO involves a customized application—for time-of-day and time-interval tasks—and the use of a smartphone’s camera function for event-based tasks (See Figure 6). A smartphone (Optus L3 II, model LG-E425f; operated on the Android™ 4.1 operating system; dimensions were 6.11 cm (W) x 10.26 cm (H) x 1.19 cm (D)) was provided to each participant for the duration of the experiment. Participants were provided with detailed instructions about how to use MEMO before they left the laboratory and were requested not to use any external aids to help remember the tasks (e.g., writing notes). In Experiment 1 there was

[15] The key element of the time-interval tasks *qua* regular in Virtual Week was that they had the same content (i.e., check lung capacity), which was a feature of time-interval tasks in MEMO, and that they occurred after a “predictable” brief lapse of time (i.e., at 2 and 4 minutes on stop clock). The variable delay in Virtual Week and MEMO for time-interval tasks differs due to practical reasons (in Virtual Week the simulation of two days is condensed into approximately 1 hour; compared to the real 24 hour time within which the MEMO is set). However, theoretically they are largely parallel in terms of the demands on monitoring that they make. That is, although these two time-interval tasks are not identical, they are nevertheless conceptually parallel in the theoretical demands they make on cognitive resources. See Naturalistic Measure section below.

a three day block just for event based tasks, and a separate three day block just for time-based tasks (PM task type blocks were counterbalanced). For the three day period that was exclusively for event-based tasks, smartphones were programmed to briefly display the selected target events each morning after participants opened a notification on the phone. For the three day period that was exclusively for time-based tasks, participants also received a notification in the morning, but in this case it was requesting the participant to select two time-of-day tasks—one in the morning (either 10am or 11am) and one in the afternoon (either 3:30pm or 4:30pm)—at which times participants were required to open the app to complete the scheduled surveys.

The smartphones were programmed to prompt participants to complete 12 event-based PM events (four on each day). Half of the events were selected by participants from a larger list of common activities and events (e.g., collect mail, feed pet, water garden). Participants were instructed to choose events that were “extremely likely” to occur during the testing days. The remaining events were selected by the research assistant. The four events for each day differed along the following dimensions: regularity; experimenter given vs self-selected (from list of options); and duration of time-window for execution (i.e., short vs long). The first two events selected were regular (i.e., required on each of the 3 days): one was set by the research assistant (i.e., lunch each day) and one was self-selected (either: brushing teeth, taking medication, or passing a daily landmark). The other two photos chosen were irregular (i.e. unique to each day): one had a short (e.g., “putting on the dishwasher”) and one had a long (e.g., “collect mail”) window of opportunity. After all events had been selected, the experimenter then programmed the chosen events into the phone.

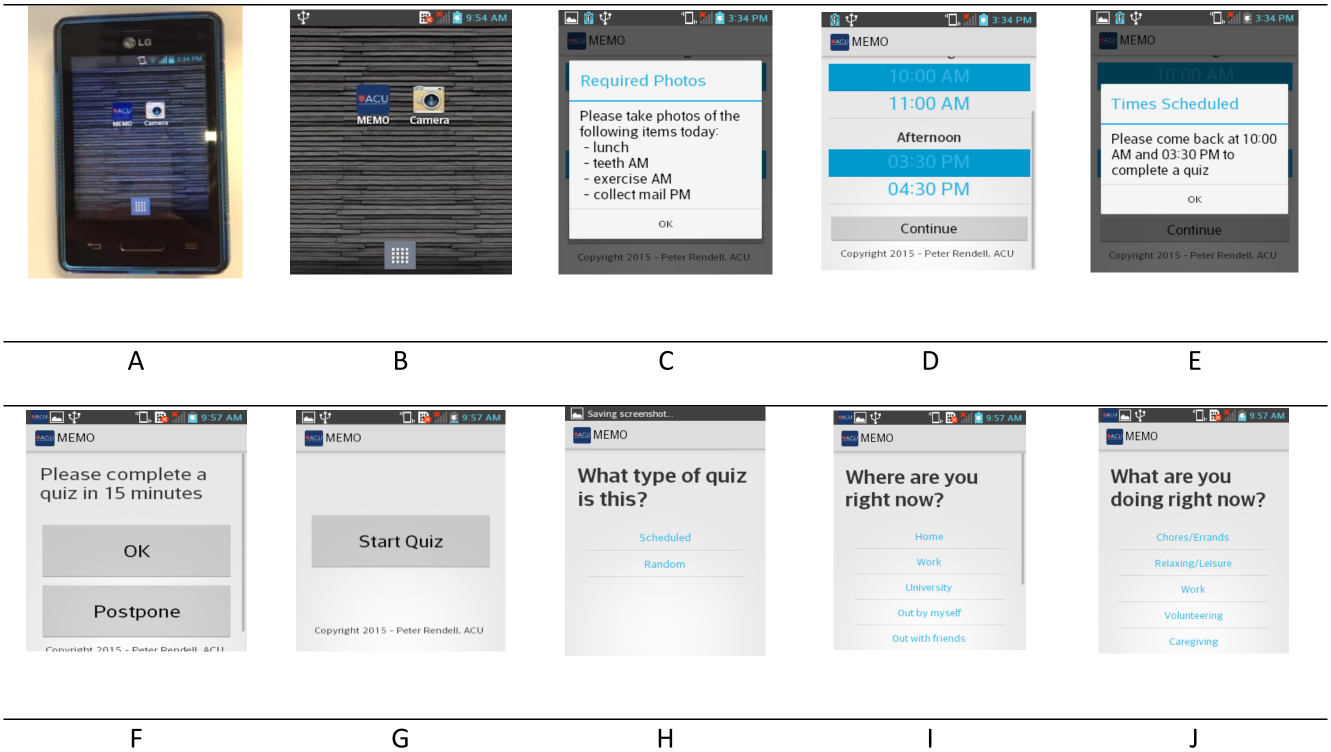

Figure 6. Screen shots from MEMO assessment of PM in naturalistic-setting. A: smartphone participants used. B: home screen showing MEMO quiz app and camera app. C: morning notification of the day's event-based tasks (require photos of events expected to occur later in the day). D: times selected for that day's time-of-day quizzes. E: final reminder of time-of-day tasks (shown immediately after "continue" button pressed in D). F: example of a random quiz notification. G: screen shown when participants remember to open app for "scheduled" (i.e., time-of-day) or "random" (i.e., interval) quiz. H: first quiz question (test of retrospective memory). I: first context question (location). J: second context question (activity). Copyright 2015 by Peter G. Rendell. Reproduced with permission.

Participants were instructed to take a photograph when the relevant events occurred, and the act of photographing the events served as the event-based PM task. The relevant list of PM tasks was displayed on the smartphone each morning of the next three days and disappeared once participant pressed a button that read "OK". Each morning, for three days following induction in the laboratory, at 9am participants received an auditory notification on the smartphone. The smartphone then displayed a list of four events that they were required to photograph that given day. Participants were instructed to only use the camera and MEMO application, such that the potential for participants trying to access previously taken photos which may act as reminders or prompts was reduced. It should be noted that sometimes participants took more than one photo of the same even when the first image was blurry. (Given that these occasional duplicate images were not deleted it is reasonable to assume that participants were not searching the phone contents for previous photos to curate images or as a reminder. Though this cannot be entirely ruled out for all participants, it is also worth noting that participants were allowed to use whichever strategies they liked to remember their tasks, which somewhat blunts the point of this potential limitation).

The time-based PM tasks were conducted in a separate three day period to the event-based PM tasks, with order counterbalanced. Both types of time-based tasks (time-of-day vs time-interval) were intermixed and administered concurrently during the three day period for time-based PM tasks. Participants were again requested not to use any external aids (e.g., alarm clocks). For the time-of-day tasks, participants again received a notification on the smartphone at 9am each day, but this time they were required to choose two times (one time-of-day in the morning and one in the afternoon) to come back and complete a brief quiz (quiz completion served as the time-of-day PM task). The times offered were 10am or 11am for the morning quiz and 3:30 or 4:30pm for the afternoon quiz. At these times participants were required to click on the MEMO icon on the home screen (the only icon displayed, apart from the camera app icon) to start their quiz. For the time-interval task, participants received two “pop-up quiz” notifications (and accompanying auditory alerts) at times unbeknown to the participant (day one: 1:30 and 5:30pm; day two: 12:30 and 6:00pm; and day three: 1:00pm and 5:30pm). These pop up notifications instructed the participant to open the MEMO app after a specified time-interval (either 10, 15 or 20 minutes) to complete another brief quiz; thus participants had to carefully monitor the time to correctly perform the time-interval task. If the participant was unable to respond to the pop up quiz notification when it was sent, then the time-interval task commenced from the time when the participant was able to acknowledge receiving the notification. In Experiment 1 the mean delay between when the six pop up quiz notifications were sent and when participant acknowledged time-interval task (i.e., began stop clock) ranged from 14–38 minutes, with a standard deviation range of 29 to 67 minutes.

The quizzes for both the time-of-day and time-interval PM tasks consisted of two multiple-choice questions: 1. “where are you at the moment?” options included “at home”, “at work”, “out by myself”, “out with friends”; and 2. “what are you doing at the moment?” options included “relaxing”, “doing chores/errands”, “working”. Experiment 2 below also included data from an additional question: “is this a scheduled (i.e., chosen by participant in the morning) or random quiz (i.e., pop up notification during the day)” to test retrospective memory.

Scoring criteria. For the time-based tasks, time-of-day tasks (scheduled each morning by the participant) were classified as correct if the MEMO application to start the quiz was opened ±5 minutes of the required time (e.g., 3:30pm). The time-interval tasks were classified as correct if the MEMO application to start the quiz was opened ± 2 minutes of the required time (e.g., opening the app after 11 minutes for a requested 10 minute time-interval would be classified as correct). Photos (event-based task) were scored using the following categories: *correct* (photo taken as intended), *missed* (no photo), *remembered forgetfulness* (participant remembered at a later time, and took a photo of a note indicating this), *reminded by another task* (photo taken one minute after another required photo), *contrived* (judged by researcher to be an unlikely photo of the true event, either due to time-stamp or appearance of event being staged), and *unable to take* *photo* (photo of note indicating event did not occur—e.g., “doctor’s appointment rescheduled”—or that it was impractical to take photo at time of event). In the current study correct photos were scored as 1 with all other categories scored as 0. The proportion of correct scores out of the total possible for all event-based (photo) tasks was the dependent variable. To assess the reliability of the scoring, a second independent researcher classified the photos taken by a subset of 25 randomly selected participants. Inter-rater reliability was computed using Siegel and Castellan (1988) variant of Cohen's kappa (1960) as the scoring was nominal. Analyses for each individual MEMO photo task (e.g., “regular” photo of lunch required on day 1) yielded kappa values between 0.84 (Day 1 photo of “regular” other—non-lunch—photo, e.g., medicine) and 1.00, indicating excellent inter-rater reliability (Hallgren, 2012).

Design. Experiment 1 used a 2 x 3 mixed factorial design. *Age group* (young, older) was the between-subjects variable and *task type* (event, time-of-day, time-interval) was the within-subjects variable. The dependent variable was mean proportion of correct responses. There were separate analyses for each setting (laboratory or naturalistic), as there is no claim that the tasks are identical across settings, rather there is conceptual parity across settings. Also, the main interest was in the pattern of results between age groups within each setting and between different task types. Partial eta-squared was used for effect sizes, with .010 classified as small, .059 as medium, and .138 as large following Cohen (1988). All statistical analyses were conducted using IBM SPSS version 22 software.

Results

Figure 7 shows the proportion of correct responses for each PM task type as a function of *age group* (young, older), *task type* (event, time-of-day, time-interval) and *setting* (laboratory, naturalistic).

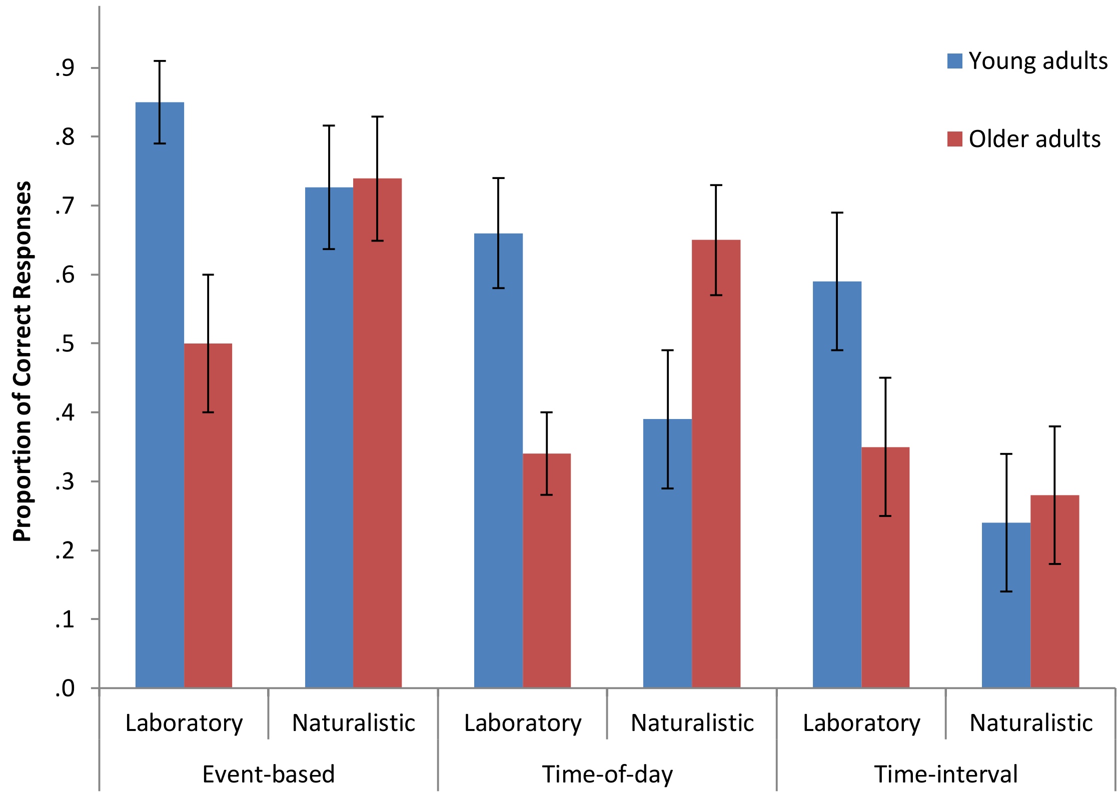

Figure 7. Mean proportion correct on the PM tasks as function of task type (event-based, time-of-day, time-interval), setting (laboratory, naturalistic), and age (young, older). Error bars depict 2 standard error of the mean (corresponding to the 95% confidence interval for the mean).

To test whether the apparently paradoxical results of previous studies may be a function of comparing disparate task types in each setting, a series of 2 (age: young, older) x 3 (task: event, time-of-day, time-interval) factorial ANOVAs were run on the data within each setting.

Laboratory Setting

In the laboratory, age group and task type did not interact, F (2, 176) = 1.27, p = .283, η_p^2 = .01, but there was a main effect for age group, F (1, 88) = 31.24, p < .001, η_p^2 = .26: with young adults (M = .69, MSE = .04) performing better than older adults (M = .39, MSE = .04). There was also a main effect of task type, F (2, 176) = 21.57, p < .001, η_p^2 = .20. Bonferroni post hoc tests showed that for all participants (across young and older adults), performance was better on event-based (M = .66, MSE = .03) compared to both time-of-day tasks (M = .49, MSE = .03; p < .001) and time-interval tasks (M = .47, MSE = .04; p < .001), while there was no difference between performance on the time-of-day and time-interval tasks (p = 1.000).

Naturalistic Setting

In the naturalistic-setting, there was a main effect for age group, F (1, 75) = 8.97, p = .004, η_p^2 = .11, and a main effect of task type, F (2, 150) = 60.15, p < .001, η_p^2 = .45, but there was also an interaction between age group and task type, F (2, 150) = 5.02, p = .008, η_p^2 = .06. Follow up analysis of the interaction revealed a difference between age groups for the time-of-day task, F (1, 75) = 17.61, p < .001, η_p^2 = .19, older (M = .65; MSE = .05) outperforming young (M = .38; MSE = .05) adults. There was no age group difference for both event (young: M = .68; MSE = .04; older: M = .75; MSE = .03), F (1, 75) = 1.82, p = .181, η_p^2 = .02, and time-interval (young: M = .27; MSE = .05; older: M = .36; MSE = .05), F (1, 75) = 1.31, p = .257, η_p^2 = .02. There was a simple main effect for task type for both young adults, F (2, 74) = 33.27, p < .001, η_p^2 = .47, and older adults F (2, 74) = 31.42, p < .001, η_p^2 =.46. For both young and older adults Bonferroni post hoc tests showed performance on event-based to be better than that on time-interval tasks (both ps < .001); while performance on time-of-day tasks was better compared to performance of time-interval tasks for both young (p = .05) and older (p < .001).

Context of correct time-based tasks performance. The brief quizzes completed by participants when performing the time-of-day (young group quiz completion range: 78–98%; older group quiz completion range: 85–95%) and time-interval tasks (young group quiz completion range: 67–90%; older group quiz completion range 60–83%), offer some insight into the context within which successful performance by young and older adults occurred. Older adults who completed the time-of-day quiz were more likely to report being at home than elsewhere (62%) compared to young adults (42%), χ^2 (1, N = 376) = 16.05, p < .001. Older and young adults reported interrupting leisure activities, i.e., "relaxing" or "eating", no more than half the time (50% vs 37%, respectively), χ^2 (1, N = 429) = 2.61, p = .106). Young adults reported their ongoing activity as "working" (44%) more frequently than older adults (10%), χ^2 (1, N = 429) = 62.96, p < .001, while older adults were more often engaged in "chores/errands" (34%) compared to young adults (13%), χ^2 (1, N = 429) = 26.51, p < .001, at the time the quiz was answered. Volunteering, caregiving, and commuting made up approximately 5% of interrupted ongoing activities for young adults and 10% for older adults, χ^2 (1, N = 429) = 3.96, p = .046.

For the time-interval task, older adults who remembered to respond were more likely to be at home (58%) compared to young adults (39%), χ^2 (1, N = 353) = 11.90, p < .001. However, a similar pattern of activities as that found for the time-of-day tasks emerged—young and older adults switched from leisure activities 45% and 50% of the time, respectively, χ^2 (1, N = 344) = 0.92, p = .338. When completing time-interval time-based tasks, young adults reported taking part in work at the time of the quiz more often (34%) than older adults (10%), χ^2 (1, N = 344) = 20.46, p < .001, whereas older adults reported task-switching from chores/errands more often (26%) than younger adults (13%), χ^2 (1, N = 344) = 10.05, p = .002.

Discussion

The design and findings of Experiment 1 provide insight into the methodological influences on the PM paradox by using conceptually similar and objectively verifiable PM task types matched across settings. The findings give both a broader and clearer picture of age-effects on PM in naturalistic-settings in contrast to previous studies that used disparate task types across settings. The apparent paradox that has emerged from previous studies, i.e. that young perform better than older adults in the lab while the reverse is true in naturalistic-settings, was replicated in our study when just focusing on the results for tasks that previous studies have used in each setting. That is, on the time-interval tasks in the laboratory (Virtual Week) older adults were inferior to young adults, whereas on the time-of-day tasks in the naturalistic-setting (MEMO) older adults were superior to young adults. The classic paradox of older adults being inferior to younger adults in laboratory-settings and superior in naturalistic-settings was also replicated when just focusing on the other typical comparison of complex event-based tasks in the laboratory (event-based tasks in Virtual Week; complex because participants had to store multiple event-based tasks at the same time), with simpler time-of-day tasks in the naturalistic-setting (simpler because participants only had to store four experimenter given time-based tasks each day; and 2 of these, i.e., the time-of-day tasks, were given each morning with a much longer delay and allowing for potential use of conjunction cues).

However, our inclusion of event-based and time-interval tasks in the naturalistic-setting (tasks which have previously almost solely been confined to the laboratory) produced novel findings that are not consistent with the PM paradox. Specifically, the time-interval task findings demonstrated that older adults did not display a universal superiority to young adults in naturalistic-settings, or that older adults performance in naturalistic-settings is always better than in the laboratory. For the event-based tasks a similar parity in performance was found between both age groups, and consistent with the great environmental support afforded by event-based cues, both groups performed better on this task type compared to the time-interval PM tasks. These findings of parity and relative performance between PM task types in a naturalistic-setting show the critical importance of using conceptually parallel PM tasks across settings in order to elucidate the main factors contributing to the age-PM paradox.

We did not explicitly compare the performance between laboratory and naturalistic-settings of either age group because the PM tasks were not strictly identical but only parallel in a conceptual sense. The tasks were similar in having relatively simple content (actions to be performed), but were nevertheless conducted in very different contexts (lab setting under experimental control vs participants' everyday life) and over different time periods (Virtual Week takes about one hour, with 10 PM tasks each virtual day that take about 10-15 minutes in real time vs 10 PM tasks over a whole day). However, our results indicate that the young adults consistently performed better in the laboratory-setting than in a naturalistic-setting. This may be due to young adults being better able to sustain attention for the relatively short time period (approximately one hour) in the lab than in the extended time (3 days) in a naturalistic-setting, or may reflect an issue with motivation. However, our results indicate that the young adults consistently performed better in the laboratory-setting than in a naturalistic-setting. This may be due to young adults being better able to sustain attention in the lab than in a naturalistic-setting, or may reflect an issue with motivation. In the naturalistic-setting, our results showed that older adults performed better than younger adults only on the task type most frequently used in this setting, time-of-day tasks. In regards to event and time-interval tasks in a naturalistic-setting, we found a parity in performance between age groups. For event-based tasks this parity has a precedent in previous laboratory studies using salient or focal event cues (Kliegel, Jäger, et al., 2008). From a parallel task perspective, the time-interval task age group parity is more surprising, given that young adults consistently outperform older adults on this task type in the laboratory. From the perspective of the age-PM paradox it is surprising that older adults did not outperform younger adults given the naturalistic-setting for this task. However, extrapolating from theoretical considerations put forward by the multiprocess framework on the relevance of self-initiated operations for some PM tasks (i.e., those low in environmental support) but not others (e.g., event-based tasks with a salient cue), it is less surprising given that the time-interval task has the least environmental support.

As a result of addressing a range of methodological shortcomings of previous studies, these findings represent a major step forward in understanding the ageing PM paradox. In particular, the choice of Virtual Week as the comparison laboratory paradigm was noteworthy. Virtual Week has the virtue of being a comprehensive measure of PM that simulates PM tasks in daily life (Phillips et al., 2008) and has been improved substantially since the first and only other study comparing Virtual Week with a measure of PM in a naturalistic-setting. It was thus possible to compare (virtual) time-of-day tasks from the laboratory with those typically used in naturalistic studies of PM, which led to a replication of the PM paradox pattern when comparing time-of-day tasks across settings. Furthermore, with the development of MEMO it was possible to compare event-based PM tasks and time-interval tasks across settings as well; the outcome of which (i.e. comparable naturalistic performance between young and older adults) indicates that more careful comparison of PM tasks across settings is needed in future studies.

Finally, by using the novel MEMO paradigm we were also able to gain contextual information (i.e., location and activity) about the nature of the ongoing tasks that young and older adults were engaged in at the time when they remembered to complete their time-based tasks. As expected, older adults were less likely to be working and were more often at home. However, it was not clear whether this necessarily meant older adults had less taxing environments or ongoing tasks compared to young adults. While older adults (45%) adults reported engaging in leisure activities more often than young adults (35%) at the time of quizzes, they were also much more likely to be involved in chores or errands (34 % to 13%). In addition, the broad category of "leisure activity" included not only rest but other activities with varying levels of demand such as "eating" and "exercise". Eating, for example, covers a wide range of activities that may vary in possible demands, such as whether or not the respondent was responsible for preparing and serving a meal or consuming one with other people (e.g., a social lunch) or alone. Thus, while it could be argued that older adults who successfully completed the time-based tasks tended to be engaged in less cognitively taxing environments (e.g., home) and tasks (e.g., leisure) compared to young adults, this conclusion needs to be made with caution in light of the lack of more detailed information regarding the exact tasks being undertaken prior to responding to the quizzes. Importantly, however, no differences between age groups were found on ongoing tasks in terms of locations or activities for the time-of-day tasks, the PM task type that is likely to be most influenced by environmental supports. These contrasting findings offer support for the relevance and importance of the distinction between the two types of time-based tasks: time-interval (high demand) and time-of-day (relatively low demand) which should be factored in when investigating the ageing PM paradox.

The largely neglected second aspect of the PM paradox (i.e., young-old outperforming old-old in the laboratory but both showing equivalent performance in naturalistic-settings) merits closer scrutiny as it is arguably *more* important to understand this second aspect of the paradox than the first given most cognitive decline occurs late in the life-span. The early findings of Rendell and Craik (2000) suggested that old-old adults were more vulnerable to PM lapses in daily life on time-interval than on event or time-of-day tasks (in the latter two PM tasks young-old and old-old adults showed comparable performance). In contrast, a more nuanced pattern of differences emerged in the laboratory, with young-old outperforming old-old on high cognitive demand (e.g., time-interval) tasks but showing comparable performance on low cognitive demand tasks (e.g., repeated event-based tasks). The second study thus continues this early work by Rendell and Craik by further probing this second aspect of the PM paradox using methodologically strengthened paradigms in both settings.

Experiment 2

As noted, Experiment 1 examined the potential influence of disparate tasks across settings as a relatively neglected factor in explaining the first aspect of the PM paradox, i.e., the apparent pattern of young adults being superior to older adults in the lab but inferior to them in naturalistic-settings. Contrary to the laboratory findings, the event-based tasks completed in a naturalistic-setting, where environmental support is usually high, showed no difference between young and older adults. The time-of-day task replicated the classic PM paradox pattern (i.e. younger adults outperforming older adults in the laboratory-setting, but older adults outperforming younger adults in the naturalistic-setting), while the time-interval task revealed a comparable vulnerability to forgetting for both young and older adults across settings. These results raise the question of whether a similar pattern of results would emerge when using parallel tasks in each setting to investigate the second key, but often overlooked, feature of the PM paradox of young-old being superior to old-old adults in the lab but there being no age differences in naturalistic-settings. Thus, in Experiment 2, we focused on differences within older adults, at a period of the life span when most cognitive decline occurs, but when more established routines and lifestyle are assumed to provide a framework of environmental support to compensate for older adults' decline in PM found in the laboratory (even on a paradigm such as Virtual Week that simulates daily life ongoing and PM tasks).

More specifically, Experiment 2 addressed the issue of whether old-old adults are more liable to PM lapses in general than young-old adults, and whether PM task types are differentially affected. It should be noted that because Experiment 2 involved comparison of young-old and old-old adults, this arguably minimized the potential influence of cohort effects that may affect extreme age comparisons (e.g. young vs older adults) and possibly contribute to young adults' inferior PM performance in naturalistic-settings. For example, the potential differing levels of motivation to complete tasks has been proposed as a cohort factor, with Ihle et al. (2012) finding that older adults were more motivated on tasks rated as low or middling in personal importance than younger adults (which might arguably be the case for experimenter-given PM tasks). Thus, Experiment 2 explicitly focused on the second aspect of the PM paradox (across the age-range of older adults a robust age-related decline in performance in the laboratory but no age differences in naturalistic-settings), reducing the potential for large cohort differences that may be present when conducting extreme age comparisons. Experiment 2 involved a larger sample of young-old and old-old adults than Experiment 1, but used the same laboratory paradigm, as well as a modified version of the MEMO. Specifically, the naturalistic event- and time-based PM tasks used in Experiment 1 were combined and administered in a single block over three days (rather than in two separate blocks of three days). Thus, the number of PM tasks required were doubled, arguably increasing the cognitive load, to a level more commensurate with that in Virtual Week, where both the event and time-based tasks are completed on each virtual day. Extending the logic and hypotheses from Experiment 1, it was hypothesized that the young-old adults would outperform the old-old adults in the laboratory, particularly on the time-based tasks, but that this pattern of differences would be eliminated or attenuated on all PM task types in the naturalistic-setting with the exception of the time-interval task, on which it was expected that the young-old would outperform the old-old in both settings, in line with the findings of Rendell and Craik (2000; Study 2).

Method

Participants

One-hundred-and-four healthy older volunteers—64 young-old (60–74 years old; 70% female) and 40 old-old (75–87 years old; 63% female) adults—were recruited through Villa Maria Catholic Homes elderly independent living units, church and community newsletters, and word-of-mouth, as part of a larger experiment on cognitive training that included the currently reported measures in its baseline phase. Participants were compensated with $50 shopping vouchers for coming into the Australian Catholic University for all testing sessions of the larger experiment. All older adults in Experiment 2 were screened using the Telephone Interview Cognitive test (de Jager et al., 2003). Those with an education adjusted cut-off score of ≤ 33 were excluded.

Materials and Procedure

The same version of Virtual Week used in Experiment 1 was used in Experiment 2[16], however, a modified version of MEMO was used, which was substantially more challenging as both event- and time-based tasks were combined over the same 3 day period (thus, there were

[16] There was one difference worth noting. As the data in Experiment 2 were collected as part of a larger cognitive training study, the version of Virtual Week also included a common measure of meta- Memory, i.e., judgment of learning questions (“how likely are you to perform this task?” with slider bar for response, 0-100%) after each task was encoded. An additional task assessing memory of the event cards encountered and the decisions made was also included—presumably encouraging more focus on the ongoing task.

eight tasks per day instead of four). The MEMO was also developed to include a “postpone” feature for pop up quizzes, which meant, if the postpone option was selected, the participant received another notification that a time-interval task was ready to be acknowledged (i.e., stop clock started) one hour later. Pop quizzes were only ever postponed once with the percentage of quizzes postponed ranging from 3.5–10.5%.

Participants were trained in and undertook the MEMO task after a session completing other cognitive measures as part of a baseline session for the larger cognitive training experiment. Approximately one week later participants completed a two day version of Virtual Week. During the training on MEMO, participants were instructed in how to open the app, perform quizzes, and take photos, and practiced doing so in the presence of a research assistant. A noteworthy difference in this induction in contrast to Experiment 1 was that participants were advised that they “could use whatever strategies they would normally use in daily life to remember a task they needed to perform”[17]. Although writing down quiz times or using external aids like alarm clocks was not explicitly encouraged, if participants explicitly asked if it was okay to use external aids, the research assistant responded by repeating that whatever strategies they would normally use or wished to use were permitted. Participants were also given a printed guide for use of the smartphone and given a support contact phone number if they had any technical difficulties. The main technical difficulty reported was that no quiz notifications or photo instructions were provided for the third day; this glitch was subsequently rectified, but a substantial number of participants consequently were only able to provide data for 2 days (young-old: n = 14; old-old: n = 10). An ANCOVA, using age group with number of days MEMO was completed as a covariate and mean proportion of correct responses as outcome

[17] This was due to Experiment 2 being part of the baseline for a larger cognitive training study in which potential changes in strategy use was of interest.

measure, revealed that there was a main effect for number of days, $F(1, 103) = 11.80$, $p = .001$. This was possibly indicative of a practice effect, with those who completed three days having a higher proportion of correct responses for the combination of all MEMO tasks[18] ($M = .88$, $SD = .17$) compared to those who completed only two days ($M = .74$, $SD = .23$). There was no main effect for age group, $p = .322$, or interaction between number of days and age group, $p = .117$. Finally, at the end of the three days, after participants returned the study smartphone, a research assistant conducted a brief post- MEMO interview over the phone, probing for what strategies participants had used and which tasks they found most challenging. As with experiment 1, data were analyzed with 2 (age) x 3 (task types) mixed ANOVAs, with separate analyses for each setting.

Results

Figure 8 shows the proportion of correct responses for each PM task type as a function of age group (young-old, old-old), task type (event, time-of-day, time-interval) and setting (laboratory, naturalistic). As with Experiment 1, in Experiment 2 to test whether the second aspect of the paradox found in previous studies may be a function of comparing disparate task types in each setting, a series of 2 (age: young, older) x 3 (task: event, time-of-day, time-interval) factorial ANOVAs were run on the data within each setting.

[18] This pattern did not change when each task type was separately analysed with an ANCOVA.

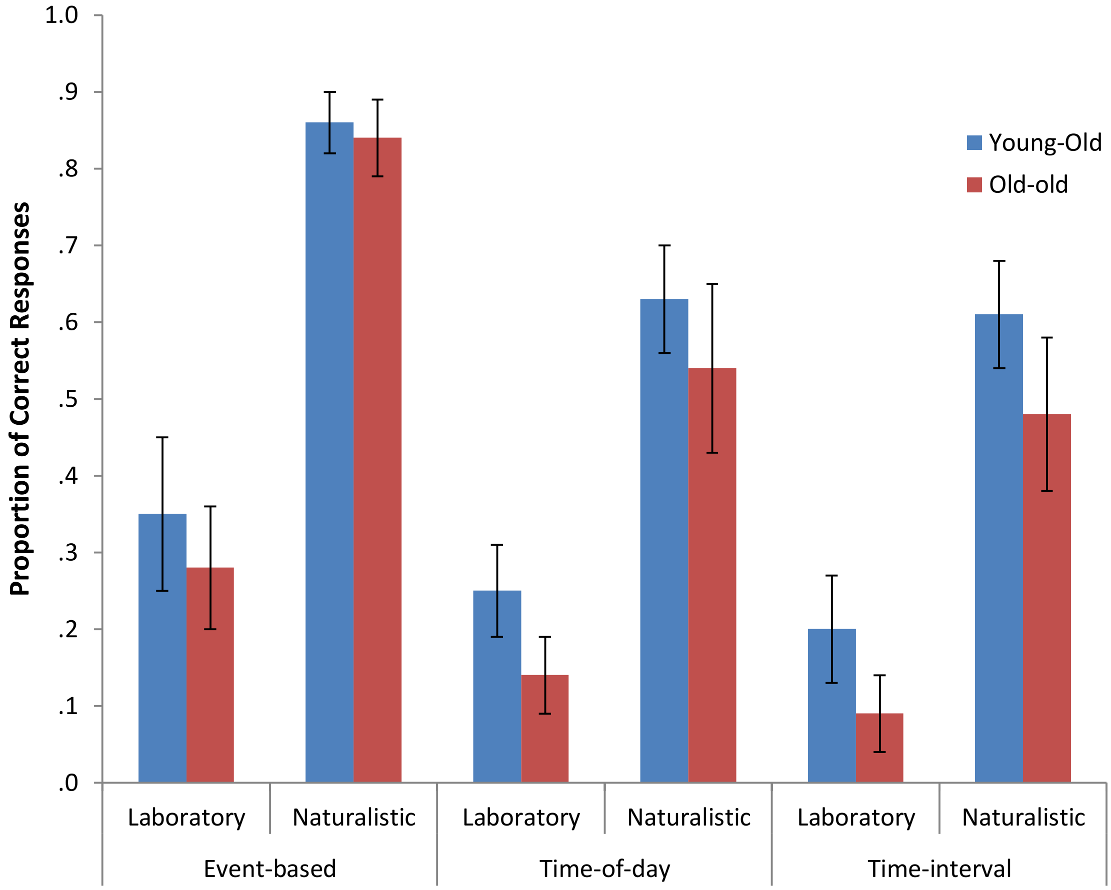

Figure 8. Mean proportion correct on the PM tasks as function of task type (event-based, time-of-day, time-interval), setting (laboratory, naturalistic), and age (young-old, old-old). Error bars depict 2 standard error of the mean (corresponding to the 95% confidence interval for the mean).

Laboratory

There was no interaction between age group and task type in the laboratory, F (2, 174) = 0.89, p = .413, η_p^2 = .01, but there was a main effect for age, F (1, 87) = 6.40, p = .013, η_p^2 = .07, with young-old (M = .26; MSE = .02) performing better than old-old adults (M = .17, MSE = .03) as shown in Figure 4. There was also a main effect of task type, F (2, 174) = 11.99, p < .001, η_p^2 = .12. Bonferroni post hoc tests showed that for all participants (across young-old and old-old groups) performance was significantly better on event-based (M = .30, MSE = .03) than time-interval (M = .14, MSE = .02) tasks (p < .001), and performance on event-based tasks was significantly better than on time-of-day tasks (*M* = .20, *MSE* = .02; $p = .027$); but there was no difference between performance time-of-day and time-interval tasks ($p = .178$).

Naturalistic

There was no main effect for age group $F(1, 93) = 3.51$, $p = .064$, $\eta_p^2 = .04$ and no interaction between age group and task type in the naturalistic-setting $F(2, 186) = 1.19$, $p = .307$, $\eta_p^2 = .01$, however there was a main effect of task type, $F(2, 186) = 14.47$, $p < .001$, $\eta_p^2 = .38$. Bonferroni post hoc tests showed that for all participants (across young-old and old-old groups) performance was significantly better on event-based tasks (*M* = .86, *MSE* = .15) than on both time-of-day tasks (*M* = .60, *MSE* = .31) and time-interval tasks (*M* = .54, *MSE* = .30; $p < .001$). There was no difference between performance on time-of-day and time-interval tasks ($p = .212$).

Context of correct time-based task performance in the naturalistic-setting. The completion of time-of-day quizzes ranged from 72–92% for the young-old group, and 63–73% for the old-old group. Responses to the time-of-day quiz showed no difference between young-old and old-old on recognition of quiz type (i.e., “is this a scheduled or random quiz?”) with both showing high accuracy (93% and 90% correct, respectively; $\chi^2(1, N = 499) = 1.25$, $p = .263$). Both young-old and old-old most frequently reported being at home at time-of-day quizzes (65% and 70%, respectively), and did not differ in any type of locations reported, *p*s > .260. When completing time-of-day quizzes young-old and old-old frequently reported being engaged in a leisure activity (i.e., “relaxing” or “eating”), 35% and 45%, respectively, $\chi^2(1, N = 499) = 0.16$, $p = .690$. The proportion of quiz responses embedded within “chores” or “work” activities for young-old (59%) compared to old-old (45%), was also not significantly different, $\chi^2(1, N = 499) = 1.70$, $p = .192$. Although the combined tasks of “commuting”, “volunteering”, and “caregiving” were infrequently reported by both young-old (6%) and old-old (10%), there was a difference, χ^2 (1, N = 499) = 6.97, p = .008, with old-old adults being engaged with these activities more often at the time of doing the quizzes.

For participants who completed the time-interval quizzes (range for young-old: 52–86%; range for old-old: 70–97%), the proportions of correct responses for the retrospective memory question were comparable between young-old (83%) and old-old (78%), χ^2 (1, N = 430) = 2.07, p = .150. Old-old adults who completed the interval quizzes were more often at home (79%) than young-old (66%), χ^2 (1, N = 446) = 8.70, p = .003. Young-old were engaged in leisure activity (57%) more frequently at the time of completing the time-interval task than old-old (43%), though this difference was not significant, χ^2 (1, N = 433) = 0.01, p = .904. Work and chores were reported as ongoing tasks for 36% of young-old and 38% of old-old, χ^2 (1, N = 430) = 0.52, p = .472. Volunteering, caregiving, and commuting combined were reported least often as ongoing tasks at the time of time-interval task completion: 7% of quizzes for young-old and 10% for old-old, χ^2 (1, N = 430) = 1.13, p = .287.

Discussion

Experiment 2 probed the second, neglected, aspect of the PM paradox, young-old superior to old-old in the laboratory but both older age groups showing equivalent performance in naturalistic-settings. Specifically, Experiment 2 continued the examination of late age PM performance differences across PM task types and settings carried out by Kliegel, Rendell, et al. (2008), and earlier by Rendell and Craik (2000) using Virtual Week and Actual Week. In both these prior studies young-old and old-old showed similar performance on time-interval tasks and regular event and time-based tasks across settings. Although it cannot be argued that, in the current experiment, the tasks administered in the laboratory are ideal for comparison with those developed for the naturalistic-setting (cf. Bailey et al., 2010), the PM tasks were conceptually parallel and in this respect represent an important departure from a long tradition of studies that have only investigated one or two—typically different—types of PM tasks in each setting.

The pattern of results for Experiment 2 in both the laboratory and naturalistic-settings were close to what the multiprocess framework would predict (McDaniel & Einstein, 2000; McDaniel et al., 2015). That is, event-based tasks that typically rely on relatively spontaneous cognitive processes were the tasks performed best by both older age groups in each setting. In the laboratory, there was a decline in performance for both age groups as environmental support declined (i.e. performance was best on event task-based tasks, followed by time-of-day tasks, and was worst on time-interval tasks). As predicted, young-old adults outperformed old-old adults in the laboratory on each task type, and age and task type did not interact.

Our findings for overall performance on the different for task types in the naturalistic-setting followed a similar pattern to those in the laboratory (i.e., performance was best on event-based, followed by time-of-day, and then time-interval PM tasks), albeit with generally better performance in the real life setting than in the laboratory-setting. This was the case for both groups of older adults, for all task types, and most notably for the time-interval task. An important finding in Experiment 2 was the reduction of the poor performance on the time-interval task in the naturalistic-setting, which was seen in the older adult group in Experiment 1. This is possibly explained by the fact that participants in Experiment 2 were permitted to use whatever strategies they wished to remember their PM tasks. (The nature and details of these potential strategies was not recorded at the time—though post hoc interviews indicated there was no difference in performance between those who used external strategies, e.g., notes or alarm clocks, and those who claimed not to—thus there is clearly an avenue for further research here.)

This finding gives cause for optimism, suggesting that older adults may be able to compensate for the detrimental effects of cognitive-aging on PM performance.

General Discussion

There is now a growing body of literature on the nuances of the PM paradox (e.g., Aberle et al., 2010; Bailey et al., 2010; Hering, Cortez, Kliegel, & Altgassen, 2014; Kliegel, Rendell, et al., 2008; Niedźwieńska & Barzykowski, 2012; Schnitzspahn et al., 2011). The current two experiments make a major addition to this literature, showing that task types need to be carefully selected in both laboratory and naturalistic-settings in order to unpack key factors contributing to both aspects of the PM paradox. Research of this nature is important, as a thorough understanding of age-effects on different types of PM tasks in laboratory and naturalistic-settings is vital to inform the provision of appropriate support to older adults in our community (Lee et al., 2018). The current experiments are the first to systematically compare the performance of young versus older, and young-old versus old-old, adults on conceptually comparable PM tasks across settings with a high degree of experimenter control in both naturalistic and laboratory contexts. The use of parallel tasks in laboratory and naturalistic-settings in Experiments 1 and 2 has clarified the age-PM paradox by showing first that the apparent pattern of age-effects differing across settings are partially a function of task type rather than setting per se and second, that age-related differences are partially an artifact of conflating two distinct types of time-based tasks: time-of-day and time-interval. Whereas the existing body of research tends to suggest that older adults are uniformly better than young adults in performing PM tasks in naturalistic-settings, our results indicated that both young and older adults PM performance is modulated by PM task types in a conceptually parallel way to in the laboratory. In particular, our results showed that, when not permitted the use of external aids, young and older adults show comparable performance on time-interval and event-based tasks in naturalistic-settings. Furthermore, the difference in performance between tasks for both older groups was in the direction expected for both settings, and thus consistent with the theory of more self-initiated operations being required as environmental support decreases (Craik, 1986; McDaniel & Einstein, 2000).

A key strength of the two experiments was the development, and use of, the naturalistic PM paradigm (MEMO), the basis for which was provided by Rendell and Craik (2000). As previously noted, MEMO addressed a number of the limitations of the Rendell and Craik (2000) naturalistic paradigm used previously (Actual Week). In particular, the access to MEMO via a mobile phone meant that it was easy and convenient for participants to record their PM task performance. In addition, the use of the mobile phone camera meant that the completion of event-based tasks was verifiable, which had not been possible in the original Rendell and Craik (2000) paradigm. The MEMO thus has wide ranging future study potential for the controlled investigation of naturalistic PM tasks that parallel those commonly used in laboratory-based studies, particularly those in the established laboratory-based measure Virtual Week, which has been used with both older adults and a large range of clinical groups (Browning et al., 2018; Foster et al., 2017; Kliegel, Rendell, et al., 2008; Mioni et al., 2017; Rendell & Henry, 2009).

The current experiments have revealed that the use of different tasks across and within studies has contributed to the contrasting age-effects, i.e. the age-PM paradox identified in the PM literature (see Henry et al., 2004). In particular, Experiment 1 has shown comparable performance for older adults on time-interval tasks in laboratory and naturalistic-settings. This type of time-based task has typically not been conceptually well matched in previous studies. For example, Schnitzspahn et al. (2011) aimed to conceptually match ongoing and PM tasks across settings by scaling up time-based tasks in the laboratory, which involved pressing an “a” key at two set times within three 12 minute blocks within a naturalistic ongoing task (i.e., watching a documentary). The apparent conceptually parallel naturalistic-setting task was sending two text messages with “a” as the content in the morning and afternoon, for three consecutive days. However, we suggest that the former (laboratory task) was a time-interval task, while the latter (naturalistic setting task) was a time-of-day task, and that each had substantially different cognitive demands.

In arguing that the theoretical understanding of PM can be enhanced by ensuring that PM task types are conceptually matched across settings, and focusing more closely in particular on subtypes of time-based PM tasks, we do not mean to imply that previous studies in the laboratory-setting give an inaccurate picture of age-effects on PM, nor that previous naturalistic studies provided a distorted or flawed picture due to lack of experimenter control over key variables. Rather it should be recognized that previous studies have differed in the types of time-based PM tasks used in each setting (e.g., Schnitzspahn et al., 2011) and that both types of time-based tasks (i.e., time-of-day and time-interval) have merit as representing *some* PM task types that occur in real life.

The conceptual framework for time-based tasks proposed here (time-interval vs time-of-day), which is based on previous empirical research (Kliegel, Rendell, et al., 2008; Rendell & Craik, 2000), provides the basis for a better matching of tasks across settings. It also provides a clear starting point for a program of research to substantially enhance our understanding of the effects of age, and of other individual differences, on PM, by considering issues such as the time period over which PM tasks are spread (up to an hour or so in laboratory tasks and over hours or days in naturalistic-setting studies). The need for further research in this areas was highlighted in

Experiment 1, which examined the same young and older participants using parallel tasks across settings, and confirmed the robust age-related differences in laboratory-setting with event-based tasks spread over approximately an hour, but also revealed that there were no age-differences when conceptually similar event-based tasks were spread over several days in the naturalistic-setting. This is an important finding as, while event-based tasks have dominated laboratory-based studies, they have rarely been incorporated into naturalistic studies.

A clear impression from the current experiments was that older adults performed much better in naturalistic-settings than in the laboratory-setting on event-based and time-of-day PM tasks. The poorer performance of older adults in the laboratory suggests that even with a paradigm such as Virtual Week, which simulates the rich contexts of daily life, there is an important gap between the conditions in the laboratory and those in naturalistic-settings (Phillips et al., 2008). In line with explanations suggested by Kliegel et al. (2007) the performance difference may also reflect the greater environmental support and potential for planning provided by the familiarity and possible salience of typically occurring, event targets available in in the naturalistic environment. The target events in the current study were pre-selected by participants from a list of options to ensure that they were very likely to occur. Being able to (strategically) select familiar target events that were “extremely likely to occur” may therefore have provided substantial environmental support for the older adults to complete these tasks relatively successfully.

Another possible explanation for older adults’ better event-based and time-of-day PM performance in naturalistic compared to their own performance in laboratory-settings is the difference in time duration. That is, spreading the tasks over a longer time period (several days in MEMO compared to around an hour for 20 PM tasks in Virtual Week), may have reduced the complexity and cognitive burden of the naturalistic PM tasks, making them more like a series of “single” targets to respond to. This would be consistent with the finding of Einstein et al. (1992) on the complexity effect, in which age-effects in the laboratory were apparent when there were multiple (i.e., four), different target events, but eliminated when there was only a single (repeated) target event.

Regarding time-based tasks, we have confirmed previous findings (e.g., Rendell & Craik, 2000; Rendell & Thomson, 1999) that older adults are very accurate in naturalistic-settings on the type of time-based tasks used in previous naturalistic PM studies (what we refer to as time-of-day tasks). This is consistent with anecdotal observations that older adults participating in research studies tend to be very reliable at attending research appointments on time; something we have observed many times even for participants who show typical age deficits on time-interval tasks in the laboratory. However, by contrast, Experiment 1 showed that older adults perform very poorly in naturalistic-settings on the type of time-based tasks that have typically only been used in laboratory studies (what we refer to as time-interval tasks). As with Actual Week in Rendell and Craik (2000), we found in Experiment 1 that the poorer performance of older adults in the laboratory (Virtual Week) on time-interval tasks was also evident for the same older adult group on another set of time-interval tasks in a naturalistic-setting (MEMO). Similar to Rendell and Craik (2000), in Experiment 1 participants were requested not to use external aids. In contrast, in Experiment 2 participants were encouraged to use any strategies they liked, including external aids (cf. Ihle et al., 2012). In Experiment 2 the generally poor performance of both age groups on time-interval tasks in the naturalistic-setting compared to all other tasks found in Experiment 1 was almost eliminated, with older adults performing at similarly high levels on the MEMO time-interval and time-of-day tasks. This provides evidence that older adults can compensate for the difficulties in monitoring short or arbitrary intervals of time.

The results of both experiments in the present study raise the possibility that the time-of-day (and to a lesser extent the event-based) tasks are apparently not as well matched across settings as the time-interval tasks. (This may be indicative of a practical constraint on any laboratory paradigm that aims to develop parallel measures to those used in naturalistic-settings). In Virtual Week, the time-of-day tasks were indicated by virtual time-of-day in the board game, and also possibly cued by the content embedded in the ongoing tasks (event cards) which involved descriptions of activities relevant to the tie of the virtual day (e.g. choosing what to eat at breakfast). However, it seems likely that the virtual time-of-day does not have the same intensity of environmental cues that would support noticing the actual time-of-day in naturalistic-settings. Moreover, the virtual time-of-day PM tasks (like the virtual event-based PM tasks) are more demanding given the tighter time window. Thus, more demanding time monitoring is probably required in Virtual Week than in MEMO for the “time-of-day” PM task type.

Converging evidence suggests that Virtual Week “time-of-day” tasks may not adequately mimic the real life demands of time-of-day tasks because Virtual Week has more variable time monitoring demands (i.e., the virtual time changes as a function of the speed at which the participant moves the token around the board, while the stop clock functions in a more, normal, linear, and inexorable fashion), as found in earlier Virtual Week studies (e.g., Rose et al., 2010). In the early versions of Virtual Week there was no virtual time-of-day clock; instead consecutive hours of the day were marked on the squares (one hour for every 8 squares). In these earlier versions, there was little difference in accuracy on time versus event-based tasks, and this may have been because moving the token past the time marked squares provided an event-based cue.

However, when the times marked on squares were replaced with a virtual clock that needed to be monitored, substantial differences emerged, with poorer performance on time versus event-based tasks (e.g., Henry, Rendell, Phillips, et al., 2012). Thus, developing a time-of-day task in the laboratory which has the same level and type of environmental support as found in naturalistic-settings is a challenge, even for a research paradigm such as Virtual Week that simulates everyday daily narratives.

There are various ways in which the method of the present experiments might be extended in future research. For example, the variable experimenter-generated ongoing tasks in Virtual Week, which have been shown to reduce older adults' performance in naturalistic-settings (Bailey et al., 2010), might be captured by embedding an event-based target within the MEMO quizzes to increase cognitive demand. For example, an additional "superfluous" question could be added to the survey as an event-based PM task cue (e.g., a question about the weather could be the cue to press a different button) The performance of this experimenter given, laboratory type task in a naturalistic-setting, could also be assessed within the context of the information gleaned about current ongoing task activity and location information provided by quizzes.

In sum, these data provide novel and important evidence about some of the neglected mechanisms contributing to the apparent age-PM paradox. They show that previous findings of superior performance by older adults in naturalistic-settings appear to be largely a function of time-of-day PM tasks being predominantly used. Additionally, they indicate that for short time-interval tasks older adults' performance was low across both settings, while younger adults' performance was higher than older adults in the laboratory-setting and at a similar low level in the naturalistic-setting. However, on the time-interval task in Experiment 2, which allowed the use of external aids, older adults performed at a particularly high level (around 70% accuracy, as compared to approximately 25% accuracy in Experiment 1 where external aids were not permitted). This pattern suggests that older adults are able to use strategies to compensate for tasks that would otherwise have high monitoring demands. Thus, the pattern that previously seemed paradoxical is no longer apparent when a distinction is made between time-interval and time-of-day tasks, and when conceptually parallel tasks are used across settings.

The present experiments also demonstrate the potential value of the MEMO as a new PM paradigm suitable for use in a naturalistic-setting, which builds on and complements the existing Virtual Week paradigm that has proved its fruitfulness in PM and aging research. The MEMO offers a high level of experimental control (e.g., by providing time-stamps for photos and quizzes) for PM tasks parallel to those already available for study in Virtual Week. Used together, Virtual Week and MEMO provide comprehensive and complementary PM paradigms that offer future research a powerful methodology for investigating various PM task features across both laboratory and naturalistic-settings. In particular the MEMO holds the promise of further illuminating and unraveling the factors driving PM performance. More generally—and in line with the diverse Virtual Week studies that emerged after that paradigm was introduced in the seminal paper by Rendell and Craik (2000)—the MEMO could readily be adapted for both a wide range of PM task dimensions and populations. The common PM deficits shown by clinical groups in the laboratory, which are assumed to be indicative of functioning in naturalistic-settings, may more directly and perhaps profitably be investigated using the MEMO, allowing both a better understanding of these groups and the possibility for tailored interventions. Further investigation of PM functioning in diverse populations, beyond the field of geropsychology, using the MEMO, may also substantially benefit from consciously designing studies with the time-based task distinction made explicit in this paper. This is a distinction that we think may prove to be as theoretically and clinically fruitful as that made in event-based tasks between focal (low in cognitive demand) and non-focal (high in cognitive demand) tasks.

Chapter 5: Empirical Study of Prospective Memory, Ageing, and Executive Functioning

5.1 Preamble

This chapter is reporting a study investigating the role of executive control in understanding age-related effects on diverse high and low demand PM tasks. As has been seen in the previous chapter, levels of environmental support for PM task types appear to interact with age-related cognitive decline (e.g., on time-interval tasks when external aids are not permitted). According to many of the theories previously reviewed in Chapter 2, this is due to more controlled or executive processes being involved on high demand (e.g., low environmental support) PM tasks, and the assumption that older adults show uniform age-related declines in these processes. However, given the variable and sometimes high level of PM evidenced by some older adults (especially with more naturalistic PM task types), there is a need to investigate individual differences in executive functioning within a healthy older adult cohort and how these differences may mediate age-effects on PM task performance, particularly on high demand PM tasks. This chapter presents such an investigation with a comprehensive range of PM measures that closely represent tasks in daily life.

5.2 Study Abstract

It has frequently been argued that executive control contributes to age-effects on older adults' prospective memory (PM). However, at present the possibility that age-related declines in executive control disrupts older adults' capacity for PM is mainly a theoretical assumption that derives from the broader cognitive ageing literature. Moreover, in the few direct empirical tests of this question that have been conducted, almost all have relied on abstract laboratory paradigms of PM. The *relative* importance of specific executive control operations (updating, shifting and inhibition) compared to other cognitive processes (such as learning, retention, and, presumably, more basic processes, such as processing speed) has not yet been clearly determined. These two types of cognitive abilities (executive and non-executive) may be parallel mediators of age-effects on PM, or they might show a more complex relationship in mediating age-effects on PM. The aim of the present study was therefore to use a wide range of PM measures (clinical, laboratory, and a new naturalistic paradigm) to investigate whether ageing indirectly affects PM via 1) individual differences in facets of executive functioning, learning, and retention (a parallel mediated model), or 2) a serially mediated model, with processing speed as the first mediator, on a range of putatively low and high demand PM tasks. Healthy older adults (n = 104; age range: 60–87 years) were administered four measures of PM: Cambridge Prospective Memory Test [CAMPROMT]; Memory for Intentions Test [MIST]; Virtual Week; and MEMO, a new naturalistic-setting paradigm with putatively low and high demand tasks. Three facets of executive functioning were also measured (using Task Switch, *N*-back and Go-no-go tasks), along with perceptual speed (using a Choice Reaction Task), and learning and retrospective memory (using the Hopkins Verbal Learning Test Revised). The results indicated that age was a significant predictor of reduced PM performance on the MIST and Virtual Week,

but that these relationships diminished in strength when executive functions were controlled for, indicating that executive functions did partially mediate the age-PM relationship. Event-based PM in MEMO was significantly predicted by learning and retention. However, executive and non-executive cognitive variables did not mediate age-effects on the CAMPROMT, which permits use of external aids. The need for circumspection in generalising the role of executive functions on diverse PM measures is recommended.

5.3 Introduction

Prospective memory (PM) is the ability to remember and execute future intentions in a timely manner in the context of other, cognitively demanding, ongoing tasks (Ellis, 1996). PM is important for independent living and PM lapses are a common complaint for older adults (Kliegel et al., 2016). Indeed, Hering et al. (2018) found that PM predicts older adults' ability to perform instrumental activities of daily living, such as remembering to pay bills on time or take medicine after a meal, and PM impairment is a potential early marker for mild cognitive impairment and dementia (Duchek et al., 2006; Jones et al., 2006; McDaniel et al., 2011; Niedźwieńska et al., 2017). However, even in healthy ageing PM function has been found to decline in later life (Kliegel, MacKinlay, et al., 2008), particularly on PM tasks that are assumed to involve higher cognitive processes such as working memory (Bisiacchi et al., 2008; Brewer et al., 2010; Einstein et al., 1992) and executive functioning (Brom & Kliegel, 2014; Henry et al., 2004). PM tasks in daily life are diverse, but can be broadly categorised into *event-based* (e.g., passing on a message to a colleague when seeing them) and *time-based* (e.g., attending a medical appointment; McDaniel & Einstein, 2007). These PM task types can also be conceptualized on a continuum from low to high in cognitive demands in terms of three criteria. First, the degree of self-initiation and putative attentional control required (Craik, 1986), particularly as a function of the type of cue which signals that the delayed intention is to be executed (Brewer et al., 2010; McDaniel & Einstein, 2000; Rendell & Craik, 2000). Thus, event-based tasks can be regarded as "high demand" if they involve a cue unrelated to cognitive processing in the ongoing task (*non-focal*), and "low demand" if they involve a salient cue or are related to cognitive processing of the ongoing task (*focal*). Second, the competing demands of the *ongoing task* that the individual is engaged in during the delay before needing to execute the PM task can influence the degree of controlled attentional processing required on PM tasks (Bisiacchi et al., 2008; d'Ydewalle et al., 1999; Kidder et al., 1997). Third, the complexity of the PM task itself, such as whether it involves a simple action response (low demand) or a series of actions (high demand) has been identified as important (Ballhausen et al., 2017; Einstein et al., 1992; Kliegel et al., 2002).

While event-based PM tasks appear to involve at least a modicum of environmental support (i.e., some external cue related to the PM task may be present that captures attention or leads to spontaneous retrieval of the intention), time-based tasks have generally been assumed to involve only self-initiated processes (i.e., executive functioning ability) and can thus be considered to fall at the higher end of the cognitive demand continuum. Consistent with this view, in their meta-analysis, Henry et al. (2004) found that older adults' PM performance showed greater decline on time-based tasks (presumed to be high demand) compared to focal event-based tasks (presumed to be low demand), but that there was no difference between time-based tasks and non-focal event-based tasks (presumed to be high demand).

However, the apparent high demand uniformity of time-based tasks may be misleading (d'Ydewalle et al., 1999; Gonneaud et al., 2011), as McDaniel and Einstein (2008, pp., italics added) note: "In time-based prospective memory tasks, the appropriate moment for executing the prospective memory intention is a particular *time-of-day* (a doctor's appointment) or *the passage* *of a particular amount of time* (taking cookies out of the oven in 10 minutes).” Time-of-day time-based PM tasks, which figure primarily in naturalistic-setting PM paradigms (Maylor, 1990; Rendell & Thomson, 1993, 1999), appear to afford relatively high environmental support in the form of a constellation of time-related cues in the environment (Rendell & Craik, 2000), such as meal times, which can be explicitly linked (i.e., made into conjunction cues) with the time when the PM task needs to be executed (thus there is some ambiguity as to whether these tasks are actually event-based; see Rose et al., 2010). Thus, time-of-day time-based tasks appear to be relatively low demand. In contrast, *time-interval* time-based tasks tend to involve “arbitrary” shorter delays (e.g., put coins in parking meter in 20 minutes time), lower environmental support, and less potential for conjunction cues, and thus appear to be high demand.

5.3.1 Theoretical Conceptualisations of Age-effects on PM

Although the specific higher cognitive functions implicated in age-related reductions in PM performance on high demand tasks—especially those of a familiar or non-abstract character (Kliegel, Rendell, et al., 2008)—have been subject to only limited direct empirical study, the relationship between age, PM, and broader cognitive processing have been extensively discussed in relation to three competing theoretical models: *the multiprocess framework* (McDaniel & Einstein, 2000; Shelton & Scullin, 2017), the *process model* of PM (Kliegel et al., 2002), and the *processing speed theory* (Salthouse, 1991, 1996). The first of these, the multiprocess framework proposes that individuals rely on either strategic monitoring or spontaneous retrieval processes to complete PM tasks (McDaniel & Einstein, 2000). The process relied on is determined by perceived task characteristics (e.g., type of cue to execute intention) and individual differences (e.g., executive attentional control), and, according to a recent development (Shelton & Scullin, 2017), involves a dynamic interplay between bottom-up (reflexive associations) and top-down (strategic monitoring) processes to meet PM task demands. The distinction between low and high demand PM tasks (and thus where age-effects would be expected to manifest most) corresponds to the distinction within the multiprocess framework of spontaneous retrieval processes (low demand) and strategic retrieval processes (high demand). In terms of the multiprocess framework, time-based tasks would generally be classified as involving strategic demands (e.g., clock checking), while event-based tasks may involve spontaneous retrieval processes when the cue is salient or focal to the ongoing task (when this is not the case strategic processes may be relied on). There is support for the multiprocess framework from research showing that older adults spontaneous retrieval processes are relatively preserved (Mullet et al., 2013) while tasks which require executive control show more age related decline (Kvavilashvili et al., 2013; Rose et al., 2010; Shelton & Scullin, 2017).

By contrast, the four phase *process model* of PM tasks (Kliegel et al., 2002), proposes that PM deficits in older adults can be accounted for by age-related declines in executive functioning. In particular, the process model proposes that executive functioning is involved when an intention is *encoded, initiated*, and *executed*. The phase between encoding and initiation is *maintenance* of the intention during an activity-filled delay, and this is argued to rely on retrospective memory ability, not executive functioning ability. Using a complex (multitasking) PM paradigm, Kliegel et al. (2002) showed that age was strongly and negatively related to performance at the intention formation phase, initiation phase, and execution phase, but not with retention, with almost all participants being able to recall the PM tasks to be performed. Furthermore, retrospective memory did not interact with scores on the executive functioning measures. However, although Kliegel et al. (2002) used a number of well-established executive functioning measures, only a single measure of PM function was used, and they conceptualised executive functioning in terms of a global construct of cognitive flexibility (cf. Salthouse et al., 2004). Consequently it remains unclear which specific aspects of executive functioning ability are involved in different phases of PM, or whether the effects identified would generalise to other measures of PM.

Finally, Salthouse's (1991) perceptual processing speed model represents a more general theory of cognitive ageing, but also has relevance to age-effects on PM specifically. A central tenet of this model is that relatively low-level information processing parameters (perceptual processing speed) is fundamental to understanding performance on higher-level cognitive operations, such as working memory and executive control, thereby suggesting a serial mediated model of age-effects on PM (Salthouse, 1996, 2000, 2014; Salthouse et al., 2004). According to one interpretation of this model, high (and possibly low) demand PM task performance is proximally affected by declines in facets of executive functioning. However, age does not directly affect these processes, but instead exerts its effects through a more fundamental cognitive ability that executive functioning depends on (perceptual processing speed). To delineate the key mechanisms of age-related change in PM it is necessary to consider this potential pathway of change as well as facets of executive functioning, learning, and retrospective memory. Thus, two ways that executive and non-executive processes may mediate age-effects are apparent, one involving parallel mediation where executive processes are of more or less relevance in conjunction with processing speed and aspects of retrospective memory; and another where processing speed accounts for all or most age related changes in PM via its influence on executive or other non-executive but related cognitive processes. There is evidence for both of these views of mediation. Azzopardi et al. (2015) found measures of attentional control mediated age-effects independently of processing speed and retrospective memory on both laboratory and naturalistic measures of PM using a large sample of older adults (n = 197). In support of a serial mediated model, statistically controlling for processing speed has been found to substantially attenuate associations of age and controlled attentional processes (Hertzog, Dixon, Hultsch, & Macdonald, 2003).

5.3.2 Facets of Executive Functioning and Age-effects on PM

Also important in understanding whether and how age-effects on PM tasks are related to executive control is to appreciate that executive control reflects a multifactorial construct, with at least three related but separable components: mental-set shifting (shifting), manipulation and updating of contents of working memory (updating), and the ability to avoid distractions and inhibit habitual responses (inhibition; Miyake et al., 2000). In the context of PM, all three components might theoretically be expected to be related. This is because *shifting* refers to the ability to maintain two or more qualitatively different goals or intentions in working memory, and control attention, redirecting it from one task or mental set to another, either strategically (e.g., clock checking) or when a cue appears (e.g. when given a direction to use a different rule of classification of stimuli). The second facet of executive functioning, *updating*, is the ability to "tag" or monitor the relevance of stimuli for a task (Cohen, Gordon, Jaudas, Hefer, & Dreisbach, 2017). Finally, *inhibition* is the ability to avoid distraction or mind-wandering (Hasher & Zacks, 1979; Maillet & Schacter, 2016) and consciously inhibit a dominant or prepotent response when required to do so (Friedman & Miyake, 2004; Miyake et al., 2000). However, while the "strategic monitoring processes" referred to in the multiprocess framework and the concept of "cognitive flexibility" highlighted in the process model suggest an important role of executive functioning in older adults' PM performance, these models fail to clarify whether all or just some specific executive components explain the variable pattern of age-effects seen on high (and low) demand PM tasks (for a similar lack of specificity with age-related decline in "cognitive resources" required for memory tasks see West & Craik, 2001).

The only two ageing studies to date to investigate all three components of executive control in relation to age-effects in PM, are by Schnitzspahn et al. (2013) and Zuber et al. (2016). In one of these studies, Zuber et al. (2016) found performance on measures of shifting, updating, and inhibition were related to multiple focal and non-focal PM measures in a large sample of young to young-old adults (20–68 years old, n = 315). Using the n-back as the ongoing task (i.e., classifying letters that appeared one at a time as either the same or different from the letter that appeared two previously in the series), one PM task (focal condition) involved a cue clearly related to the cognitive processing of the ongoing task (i.e., pressing a different key on the computer when an A or D appeared), while the other PM task (non-focal condition) involved a cue unrelated to the cognitive processing involved in the ongoing task (i.e., pressing a different key on the computer when the letter was surrounded by a particular coloured frame). Contrary to the predictions of the multiprocess framework, for the focal condition measures of updating and inhibition showed small (r = .12 to .19) but significant associations with PM performance. A similar pattern, with the addition of shifting, was found in the non-focal condition. The other study by Schnitzspahn et al. (2013) used only a high demand, non-focal PM task, embedded in an abstract laboratory paradigm similar to Zuber et al. (2016). The study by Schnitzspahn et al. (2013) found shifting and inhibition, but not updating, partially explained the association between age and PM. However, an important limitation of these two studies was the use of only one, abstract, laboratory-based measures of PM function, which fails to closely represent the types of PM task encountered most frequently in everyday life. Further work is therefore needed that directly establishes the role of broader age-related decline in executive control operations to understanding age-effects in PM function.

Studies that have not simultaneously compared all three facets of executive functioning have produced equivocal support for the mediating role of executive function in the association between age and PM (Kliegel, McDaniel, et al., 2008). For example, using a large (n = 330) lifespan sample (18 to 89 years), Salthouse et al. (2004) found only a small (non-significant) relationship between executive functioning and age (r = -.13). However, for a smaller subsample of older adults aged between 50-89 who correctly recalled instructions for the PM tasks given (n = 71) individual differences in executive functioning was significantly correlated with PM (r = .49). However, age remained significantly related to PM even after partialling out variance accounted for by executive control (r = .33). Salthouse et al. (2004) cautioned against singling out executive functioning as being the main variable in accounting for age-effects on PM, noting that the relationship of PM with fluid intelligence (r = .63) and processing speed (r = .63) with PM was comparably large.

In terms of non-laboratory PM outcome measures, Kamat et al. (2014) found a global executive functioning measure was significantly associated with time (high demand; r = .31) but not event-based (low demand; r = .20) tasks for older adults using the MIST (Raskin et al., 2010), a clinical measure of PM. In contrast, Wilson et al. (2005) using the CAMPROMT—another clinical measure similar to the MIST in respect to including an even number of event and time-based tasks—found a significant association of a global measure of executive functioning with event-based tasks (r = .50) but not with time-based tasks (r = .26). A distinction between the MIST and the CAMPROMT that might shed light on this latter finding is that only in the CAMPROMT are participants allowed to take notes, in which case they might record the times at which time-based tasks are to be executed, thus effectively off-loading the cognitive demands of this task (Gilbert, 2015) and attenuating the relationship with executive functioning. Careful attention to the types of event-based tasks in the CAMPROMT may also reveal some of these involve non-focal cues or unfamiliar associations between cue-target action pairs, which would account for the high correlation with executive functioning reported by Wilson et al. (2005).

Earlier work by Salthouse et al. (1991) reinforces the need for further research to investigate more complex mediated models of age-effects on PM. At present, two main models of mediated age-effects on PM have been investigated, namely, parallel (in which multiple aspects of cognitive ability independently mediate the relationship between age and PM performance) and serial (in which one or more mediators occur in a "chain of indirect effects"; schematically, the main predictor X effects W which effects Z which most proximally effects outcome variable Y). To date, few studies have investigated whether there are parallel mediated age-effects. Cherry and Lecompte (1999) found a small but significant effect of updating partially mediating the relationship between age and PM. Similarly, Rose et al. (2010) found that updating was associated with PM performance on high demand event-based PM tasks. However, as previously noted, Schnitzspahn et al. (2013) found updating did not account for the association between age and performance on a high demand event-based PM task, potentially indicative of the different PM paradigms used.

Gonneaud et al. (2011) investigated a wide range of variables that potentially mediate the relationship between age and event-based and time-based PM tasks. They found that retrospective memory and learning (binding items of information together or forming strong associations) fully mediated age-effects on PM performance on both low and high demand event-based PM tasks (but not on time-based tasks). In contrast, no evidence was found to support the mediational role of shifting, updating or inhibition on the association between PM and age. Perhaps the most wide ranging study of executive functioning and retrospective memory mediating age-effects within an older adult cohort on diverse PM tasks was carried out by Azzopardi et al. (2015) using structural equation modelling. In that study two traditional laboratory event and time-based tasks were used, as well as a letter mailing task (five letters to be mailed back on specific days) and an experimenter-controlled event-based task (a pre-arranged phone call made by the experimenter was the cue for participants to perform a PM task; i.e., ask when the next phone appointment is). In estimating a single PM factor, the loadings for the laboratory tasks ranged between 0.528 (event-based) and 0.616 (time-based), while the naturalistic PM measures were less related, letter task 0.178 (marginally significant $p = .055$) and experimenter controlled phone call cued task 0.328. Azzopardi et al. (2015) found support for a moderated mediated model with executive functioning (measured by the Trail Making Test b) interacting with age in predicting performance on PM (defined as an overall score across all PM task types). Azzopardi and colleagues also measured processing speed and retrospective memory which were found not to significantly mediate the age-PM relationship.

5.4 The Current Study

There has been qualified support for the predictions made by the multiprocess framework, the four process model, and the processing speed model, that the relationship between PM and age is at least partially mediated by executive functioning ability on more demanding PM tasks (i.e., non-focal, time-based, complex or multiple PM tasks being simultaneously maintained), while low demand PM tasks (i.e., focal event-based tasks) show minimal or no age-effects as they are relatively immune to declines in controlled processing. However, until recently, many studies that have investigated the relationship between executive functioning and PM performance have used global measures of executive functioning (McDaniel, Glisky, Guynn, & Routhieaux, 1999; Salthouse et al., 2004), while those that have measured different facets of executive functioning ability focused on only one or two PM task types using abstract laboratory paradigms (Schnitzspahn et al., 2013; Zuber et al., 2016). Although using one or two well controlled laboratory measures of PM may help to isolate the role of executive functioning on these task types (cf. Marsh & Hicks, 1998) it also raises a question as to the generalisability of results to other common PM tasks (e.g., the PM tasks used in clinical measures and those used in naturalistic-settings), particularly those with less tight time constraints. There is also still the open question of whether older adults' individual differences on facets of executive functioning—which may proximally account for PM performance on some (if not all) PM tasks—is a function of individual differences in perceptual processing speed (Salthouse, 1996) for these more naturalistic PM tasks.

To progress the understanding of which facets of executive functioning may mediate the relationship between age and PM, the present study included a comprehensive set of measures of PM to represent putatively low and high demand PM tasks. In particular, two clinical measures of PM (MIST and CAMPROMT), a laboratory paradigm of PM with multiple PM targets (Virtual Week, which has a high number of PM trials as opposed to the two or three normally found in the classic laboratory paradigm), and a new naturalistic-setting PM paradigm (MEMO; Randall, 2016) were used. This broad mix of PM measures is important as most previous research findings have come from event-based tasks in abstract laboratory paradigms (Henry et al., 2004), while few naturalistic-setting investigations of age-effects being mediated by executive functioning ability on high demand tasks have been carried out (Azzopardi et al., 2015). Furthermore, there is mixed evidence on whether the relationships between age and PM on putatively high demand time-based PM tasks of the type found in clinical measures are mediated by executive functioning ability (Kamat et al., 2014; Wilson et al., 2005). In the present study, the three facets of executive function proposed by Miyake et al (2000) were examined. To address previously identified issues relating to the measurement of mental set shifting (i.e. the possibility that some participants may trade-off time for accuracy), a rate residual metric was calculated (Hughes et al., 2014). Given there is some evidence and theory that processing speed is more fundamental than executive functioning ability (d'Ydewalle et al., 1999; Gonneaud et al., 2011; Salthouse, 1996; Salthouse et al., 2004) in accounting for age related variability on PM task types, the present study includes a measure of processing speed, in addition to measures of facets of executive functioning, to investigate this possibility. Similarly, given Gonneaud et al. (2011) finding that retrospective memory mediated older adults performance on event-based tasks, regardless of whether they were high or low demand, a measure of learning and retention (retrospective memory) was also included in the present study.

The aim of this study was to examine the role and relative importance of facets of executive functioning ability in the association between age and PM, along with processing speed and retrospective memory, using a wide range of PM tasks differing in environmental support (i.e., low and high demand). Consistent with the multiprocess framework (McDaniel & Einstein, 2000; Shelton & Scullin, 2017) and the process model (Kliegel, MacKinlay, et al., 2008; Kliegel et al., 2002) and based on previous findings in the laboratory (Henry et al., 2004; Rose et al., 2010; Schnitzspahn et al., 2013; Zuber et al., 2016), it was hypothesized that age would influence PM indirectly through facets of executive functioning ability. However, as per the multiprocess framework we expected these indirect effects to be largest for high demand PM tasks (i.e, generally larger for time-based than for event-based tasks; and larger for PM tasks where note taking was prohibited than for PM tasks where note taking was permitted). Furthermore, it was expected that both processing speed and aspects of retrospective memory would show a similar, mediating relationship for the age-PM association only on high demand PM tasks. Given equivocal results of past studies, we were unable to postulate hypotheses about which specific facets of executive functioning would have the most robust effects on the association between PM and age for different types of PM tasks. Our results on how components of executive and non-executive cognitive processes relate to both low and high demand PM tasks has the potential to yield a better understanding of age-related decline in PM performance.

5.5 Method

5.5.1 Participants

One-hundred-and-four healthy older adults volunteered to participate, $M_{\text{age}} = 72.65$; SD = 6.36 (range 60–87; n = 70 female). Participants were recruited through Villa Maria Catholic Homes, elderly independent living units, church and community newsletters, and word-of-mouth, as part of a larger experiment on cognitive training that included the currently reported measures in the baseline phase. Participants all had English as their first language; the mean number of years of education was 14.92 (SD = 3.98; range: 7–31); and the mean estimated verbal IQ for the sample, using the National Adult Reading Test (Nelson & Willison, 1991), was 113.66 (SD = 7.80; range: 87–127). All participants were screened using the Telephone Interview for Cognitive Status modified test (TICS-M; de Jager et al., 2003), a widely used measures to screen for cognitive impairment among older adults, with an education adjusted cut-off score of ≥33 (Lacruz et al., 2013). The mean score for TICS-M, adjusted for education, was 37.83 (SD = 2.89; range: 34–45). Participants were compensated with $50 shopping vouchers for coming into the Australian Catholic University for all testing sessions of the larger experiment. The research for the parent study, from which the present study utilises the baseline data collected primarily by the first author, was approval by the Human Research Ethics Committee of the Australian Catholic University (2015-259H).

5.5.2 Materials

PM measures.

The Memory for Intentions Test (MIST; Raskin et al., 2010). The MIST is a clinical measure of PM with demonstrated construct validity with healthy older adults (Kamat et al., 2014). Standard administration time for the MIST is 25-30 minutes. There are a total of 8 PM trials that both differ and overlap on various dimensions, forming six subscales: 2-minute time delay; 15-minute time delay; time cue (e.g., “in 15 minutes tell me to check my mail”); event cue (e.g., “when I hand you a pen write the date on word search paper”); verbal response (e.g., participant needs to tell examiner: “check your mail”); and action response (e.g., participant needs to write down name of a person to contact in an emergency). The six subscales scores are calculated by summing the PM raw scores on combinations of 4 different trials. Scoring for event based tasks is either 2 (correct) or 0 (incorrect) points, whereas for time-based tasks it is either 0, 1, or 2 points, depending on whether the participant remembers they need to do something, but not the content of the task, or is more than 1 minute late or between 1-2 minutes early (both scenarios 1 point). Higher scores indicate better PM performance, with possible range for each subscale 0–8, and the total PM score 0–48 (i.e., summing scores from all subscales). A digital clock is used for participants and examiners to keep track of time and the ongoing task is a word search puzzle. The measure was administered following the guidelines in the administration manual (Raskin et al., 2010).

Virtual Week brief version (VW). VW is a well-established PM research paradigm (Rendell & Craik, 2000; Rendell & Henry, 2009). In the brief two day version of VW, plausible, simulated PM tasks are given in a computerised board game format. Administration time with older adults varies, but generally takes between 30-45 minutes. Participants roll a die to move a token around squares on the board, with each 2 squares representing 15 virtual minutes, and each full lap of the board representing a day. Before they begin moving around the board, participants are told they must complete a range of PM tasks at different points during the day. Specifically, six health tasks for both assessment days (e.g. take antibiotics at breakfast) and a unique event (e.g., pick up dry cleaning when shopping) and time-based task (e.g., attend hair dresser appointment at 1pm) for each day. While moving the token around the board event cards are picked up and read immediately after passing squares on the board marked with an E (for event). Another specific event and time-based task for the “day” is given after picking up one of the event cards while moving around the board. The event cards help construct a plausible narrative of daily activities that forms the ongoing task. The virtual time of day is also displayed in the center of the screen beneath the die, while above the die there is a stop clock for *time-interval* PM tasks (e.g., at 2 minutes on stop clock check blood pressure). The PM tasks are completed by clicking on a *perform task* icon. Participants are given a list of possible tasks and click on the one they believe it is necessary to carry out at that time/event.

There are a total of 20 PM trials over two virtual days (8 time/“virtual time”; 8 event; and 4 real time-interval). Prior to these two test days there is a practice day where participants learn the features of the game. Each virtual day has two regular and two irregular event-based and time-of-day PM tasks, and two regular time-interval PM tasks. Thus there are five subscales: regular event, regular time, irregular event, irregular time, and time-interval. Correct responses are those that occur after the specified event or time square has been just landed on or passed and before rolling the die for the next move. Time-check tasks are correct either at the exact time specified (e.g., 2 minutes) or within 10 seconds after that time. Scoring is carried out by taking the proportion correct over the two days for the total number of tasks and for each subscale (i.e., calculated by dividing the actual number correct by the total number of possible correct tasks), thus higher scores indicate better PM performance.

The Cambridge Prospective Memory Test (CAMPROMT; Wilson et al., 2005). The CAMPROMT is a clinical measure of PM with an administration time of 25-30 minutes. It was administered in accordance with the instructions in Wilson et al. (2005). It includes a total of six PM tasks: three time-based (e.g., "in seven minutes" time I would like you to stop whichever task you are on and change to another task") and three event-based PM tasks (e.g., "show and name these 5 objects, when I tell you 'the test is over'"). The required retention of PM tasks ranges from 7-20 minutes, with delays balanced across task type. Both a kitchen countdown timer (beginning at 20 minutes and counting down) and an analogue clock (set at 12pm and counting up) are used as reference points for time-based tasks. The ongoing/distractor task involves a range of puzzles and trivia questions. Note taking is permitted during testing. Prompts are given by the examiner if a participant does not respond at the required time/event or carries out wrong task. If the response after the first prompt is still incorrect, a second prompt is given.

The test has two subscales; possible scores for subscales of time or event-based tasks are 0-18, with possible scores for total test 0-36, with higher scores indicating better PM performance.

MEMO. The MEMO was designed to measure event and time-based PM in daily life using a smart phone app. Participants were provided with phones with the MEMO app for three days.

Time-based measures involve two scheduled quizzes each day at set times chosen by participant (one in the morning and one in the afternoon), and two random quizzes which require the participant to open the app when they receive an auditory notification and then again after a specific delay (range 10–20 minutes). Event based measures involve four photos to be taken each day when participants encounter particular events. Two of these photos are for the same event each day; one assigned by the experimenter (take photo of lunch), the other selected from a list of three options (brushing teeth, taking medication/vitamins or passing a daily landmark). The other two photos are chosen from a list of likely events and are unique tasks to each day (e.g., feeding pet). Participants are reminded of the 4 event-based tasks at the start of each day with a notification from the app at 8am, and are also prompted to select a (scheduled) morning and afternoon quiz at that time (either 10am or 11am in the morning; and either 3:30pm or 4:30pm in the afternoon).

Event-based tasks (i.e., photos) are scored categorically as *correct*, *missed*, or *remembered forgetfulness* (i.e., if participant notifies researcher that they remembered after the task was due). Time-based tasks are measured as the time deviations between the time the quiz is responded to and the time the quiz is due (e.g., if the app notification specifies “come back to do quiz in 15 minutes” and the participant responds to quiz in 16 minutes time, then it would count as a 1 minute deviation). Responses for time-of-day (scheduled) quizzes are classified as *correct* if they occur ± 5 minutes of the target time, and correct for time-interval tasks if they occur ± 2 minutes of the target time. The event and time-based subscale scores are calculated as the *proportion of correct responses* out of the total number of event/time-based tasks.

Measure of Learning and Retrospective memory

The Hopkins Verbal Learning Test Revised (HVLT-R). The HVLT-R is a measure of verbal encoding, learning and retention, and thus helps control for the retrospective component of PM tasks (i.e., the content of a task that needs to be encoded and recalled at the correct time). The HVLT-R involves participants listening to and memorising a list of 12 words, which are read out at approximately two second intervals. After the list is read, participants use free recall to tell the experimenter as many words as they can remember in any order. The list is then read a second and third time following the same procedure. Finally, after a 20-25 minute interval, participants are given a surprise delayed recall test for the list of words and a forced choice (i.e., yes-no) recognition test with twelve semantically related (lures) and unrelated words. Scoring involves summing the correct responses on trials 1, 2, and 3 (list readings) for a Total Recall raw score. A percentage score is given for the Retention raw score, which is the Delayed Recall score divided by the highest number of words recalled from any of the first three recall trials, multiplied by a hundred.

Executive Functioning (EF) measures.

All EF measures and the Choice Reaction Task for processing speed (see control measures subsection below) were administered on a DELL Latitude E7440 laptop computer using E-prime software (Psychology Software Tools, Pittsburgh, PA).

Task-switch task (TST). The TST is a speeded task that involves classifying stimuli according to one of two rules as quickly as possible. The stimuli presented are bivalent; meaning that each stimulus could be responded to using either rule of classification. However, integral to this task, participants are requested only to use the rule of classification referred to in the question presented at the top of the screen on each trial (i.e., either "greater or less than?" or "odd or even?"). Three blocks were used with 40 trials in each pure block (i.e., only “greater or less than?” or “odd or even?” trials, presented first and second, respectively). The third, mixed block had 80 trials, with an equal number of each classification rule, and 20 repeat trials (same rule following a trial), and 40 switch trials (different rule following a trial). The rate residual score, with higher residuals from expected scores indicating better performance (Hughes et al., 2014), was calculated for each participant. This single metric has been found to have good validity and higher reliability when some participants have similar response latencies but different rates of commission errors on switch trials (Hughes et al., 2014).

N-back. The *n*-back is a widely used measure of working memory (e.g., Salthouse et al., 2004). It involves classifying a series of letters presented briefly one at a time on the screen as the same or different to a particular letter preceding it in the series by clicking on either the right (not a target/different) or left (target/same) mouse button. In the present study this was the letter shown two previous to the current letter (i.e., 2-back). For example, in a series beginning: A, P, W, P, … the first three letters would be classified by a right click (i.e. non-targets or different), and the fourth with a left click (i.e., a target or the same), because it matches with the letter two back (i.e. P). The scoring is the proportion of correct categorisations over all trials.

Go-no go. The “go-no go” is a measure of inhibition of a primed or pre-potent response. A series of letters is presented briefly one at a time on the screen and participants need to press a key on the computer keyboard (i.e., the down arrow) as quickly as possible for all letters (180 or 75% of all trials collapsing across blocks) except the letter X (appearing on 60 or 25% of all trials collapsing across blocks), for which no response is to be made. The dependent variable was the proportion of commission errors on all no go trials (collapsed across blocks).

Processing Speed Measure.

Choice Reaction Task (CRT). The CRT was used to measure perceptual processing speed. In this task there are four blocks, with two simple stimulus presentations (a blue and a red square) alternating pseudo-randomly across trials within each block. Using the two buttons on a computer mouse participants had to click on the left side if they saw a red square and on the right side if they saw a blue square as quickly as possible without anticipating what the next color would be. The dependent measure for this task was mean latencies for all correctly classified stimuli (accuracy, i.e., pressing the correct button for each colour presentation, ranged between 90-100% for all participants).

5.5.3 Procedure

Participants were tested individually in two sessions of 2-3 hours in a laboratory at the Australian Catholic University in Melbourne (or at the participant's home if mobility was an issue). In Session One, the MIST was administered and the induction for the MEMO was given (with testing occurring over the following 3 days). The measures of executive functioning, learning, retention, and processing speed, were also administered. In Session Two, the remaining two tests of PM, Virtual Week and the CAMPROMT were administered (see subsections below for details) prior to an induction into the cognitive training parent study. To minimize fatigue, breaks were provided as needed throughout the testing sessions in the laboratory. All tests were presented in a fixed order.

Statistical Analyses.

Descriptive statistics (means and standard deviations and correlations) were computed for all measures used in the study. The parallel mediation model of the associations between age, cognitive measures, and PM were investigated with two stage hierarchical linear regressions. The serial mediated models were assessed using path analysis. Descriptive statistics and regression analyses were conducted with SPSS version 25. The parallel mediated models were tested using the process macro for SPSS (Hayes, 2018), and the serial mediated models of age-effects on PM were tested using Mplus Version 8 (Muthen & Muthen, 2017). Alpha level of 0.05 was used to interpret the statistical significance of inferential tests.

5.6 Results

5.6.1 Preprocessing

Thirteen participants only did the first baseline session (withdrew from parent study, mostly citing that the parent project was too time demanding), and two participants had an extreme hiatus between testing sessions (255 and 258 days) due to a range of non-health related personal issues. Only the data from the first baseline session was included for these two participants. For the remaining 89 participants, all of whom completed Virtual Week, the average number of days delay between the two baseline sessions was 15 days (*SD* = 9 days). (Note that the MEMO was commenced the day after the first baseline, so after three days of MEMO was completed the delay until the next session, on average, was 12 days). From these 89 participants only 79 also completed the CAMPROMT; the first 6 participants were not administered CAMPROMT at the beginning of parent study due to protocol at the time, and 4 participants' data was missing due to a clerical error. There was missing data for 4 participants for the MIST also due to a clerical error. Finally, due to a technical issue in MEMO (which is administered via a smartphone) data for some participants (*n* = 24; 23%) was only available for two of the three days. Exploratory analysis showed that PM performance was better for the participants who completed all three days (perhaps due to practice effects), and thus participants for whom only two days of data was available were omitted from the MEMO analyses.

Participants whose accuracy performance on the TST was below 50% were classed as having missing data for the TST (*n* = 6), as this suggested they did not understand the task, but were included in all other analyses. One participant was removed from analyses for the Go-no go task due to an excessive number of large deviations from their mean latency on go trials (56% of go trial latencies were more than 3 standard deviations above their mean latency, suggesting either a misunderstanding of the task or lack of motivation to comply with instructions to respond as quickly as possible on go trials).

5.6.2 Descriptive Statistics

The descriptive statistics for the PM and EF with other cognitive measures are shown in Table 5.1. It can be seen from Table 5.1 that participants performed reasonably well on almost all PM measures, as reflected by the negative skew which indicates that the majority of participants obtained high scores on the respective measures. Where skewness exceeded ±1, the variables were initially transformed to improve normality for inferential tests, however comparison of analyses with untransformed and transformed variables showed no substantial differences, except for the VW time-interval measure (which was positively skewed) and the MEMO event-based measures (which was negatively skewed). Therefore, only for these two measures were the transformation retained (all other analyses were made using untransformed variables). Performance was generally low on the Virtual Week task, with means for proportion correct below .5. The reverse (all means above .5) was the case for the average performance on the MEMO proportion correct.

Table 5.1

Descriptive Statistics for PM Measures

PM Measures	*N*	Min	Max	*M*	*SD*	Rel.	Skew.	Kurt.
CAMP Tot	79	13	36	25.95	5.29	.53	-0.46	-0.28
CAMP EB	79	2	18	12.90	2.85	.28	-0.84	1.51
CAMP TB	79	4	18	13.10	3.69	.43	-0.42	0.42
MIST Tot	100	6	48	31.59	8.38	.43	-0.49	-0.16
MIST EB	100	0	8	5.95	1.89	.31	-0.67	-0.07
MIST TB	100	0	8	4.78	1.74	.17	-0.27	0.04
VW Tot	89	0	.70	.25	.16	.69	0.64	-0.13
VW EB	89	0	1.00	.31	.27	.65	0.81	-0.17
VW TB	89	0	.88	.21	.21	.40	0.88	0.12
VW TI	89	0	1.00	.16	.23	.33	1.22[b]	0.88[b]
MEMO Tot	92	.38	1.00	.76	.17	N/A[a]	-0.47	-0.77
MEMO EB	94	.25	1.00	.85	.17	.69	-0.77[b]	-0.23[b]
MEMO TB	100	0	1.00	.72	.24	.27	-0.58	-0.38
MEMO TI	96	0	1.00	.71	.26	.59	-0.72	-0.21

CAMP Tot: CAMPROMT total score; CAMP EB: CAMPROMT event-based score; CAMP TB: CAMPROMT time-based score; MIST Tot: MIST total score; MIST EB: MIST event-based score; MIST TB: MIST time-based score; VW Tot: Virtual Week total proportion correct score; VW EB: Virtual Week total event-based task proportion correct score; VW TB: Virtual Week total time-based task proportion correct score; VW TI: Virtual Week total time-interval task proportion correct score; MEMO Tot: MEMO total proportion correct score; MEMO EB: MEMO total event-based proportion correct score; MEMO TB: MEMO total time-based task proportion correct score; MEMO TI: MEMO total time-interval task proportion correct score. Rel: Cronbach's alpha.

[a] Due to pattern of missing data it was not possible to calculate the reliability of the MEMO total score / reliability could not be estimated from the data. [b] The values for the skewness and kurtosis correspond to the transformed variables used in hierarchical regressions analyses. Logarithm 10 transformation was carried out for Virtual Week time-interval proportion correct, as many participants were at floor (*n* = 52); a reflect and square root transformation was carried out for the MEMO event-based proportion correct score as most participants were at or near ceiling level (at ceiling *n* = 19). [c]Variable not transformed; analyses with and without transformations for this variable (including with and without transformation of the two PM measures that were transformed) did not make a significant difference to outcome measures.

Descriptive statistics for the measures of executive functioning are shown in Table 5.2. The TST scores (rate residuals), with higher residuals indicating better shifting, and Go-no go scores, with higher scores indicating lower inhibition (i.e., increased number of commission errors) were both positively skewed (See Table 5.2). However, analyses with and without log transformed variables revealed no significant difference in PM outcome measure, and thus the original scoring was kept to facilitate interpretation. N-back response latencies on correct trials were within a reasonable time range (i.e., no latencies less than 300 milliseconds).

Table 5.2

Descriptive Statistics for Executive Functioning Measures

	N	Min	Max	*M*	*SD*	Rel.	Skew.	Kurt.
TST	87	-0.33	0.74	0.00	0.16	—	1.32[a]	4.05
N-back	88	.37	.93	.78	.10	—	-1.56	3.30
Go-no go	89	0	25	6.16	4.33	—	1.47[a]	3.56

TST: Task switch task; N-back score is for proportion of hits; Go-no go score is total number of commission errors over all blocks; [a]Variable not transformed, analyses with and without transformations for this variable (including with and without transformation of the two PM measures that were transformed) did not make a significant difference to outcome measures.

Table 5.3 presents the descriptive variables for the other cognitive process measures. The range for processing speed was acceptable (no mean latencies were less than 300ms). Average retrospective memory for both highest immediate recall of items in the first three trials and the delayed recall trials of HVLT-R was relatively high, reflective of a cognitively healthy sample.

Table 5.3

Descriptive Statistics for Processing Speed, Learning and Retention Measures

	N	Min	Max	*M*	*SD*	Rel.	Skew.	Kurt.
CRT	88	379	997	526	105	—	2.02[a]	5.981
HVLT Tot Rec	92	6	12	10	1.51	—	-0.52	-0.140
HVLT Ret	92	33	100	89	15	—	-1.62[a]	2.354

CRT: Choice reaction task, response latencies in milliseconds for all trials; HVLT-Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT Ret: HVLT-Ret: Hopkins Verbal Learning Test Revised retention percentage score. [a]Variable not transformed; analyses with and without transformations for this variable (including with and without transformation of the two PM measures that were transformed) did not make a significant difference to outcome measures.

A correlation matrix of all variables is shown in Table 5.4. Age was correlated in the expected direction (i.e. greater age was associated with worse performance) with the majority of measures. Weak positive correlations, or no correlations, were found between age and MEMO, and age and inhibition. Longer average processing speed times was associated with lower MIST total score and time-based score, while processing speed appeared to be unrelated to event-based scores on the MIST and the other PM outcome measures. The PM total and subscale measures generally correlated with one another in theoretically meaningful directions. The measures of facets of executive functioning ability did not correlate with each other, though the switching (TST) and updating (n-back) tasks correlated with processing speed. The measure of inhibition (Go no go) was negatively correlated with the proportion correct on event-based tasks in the MEMO (i.e., better inhibition—less commission errors—was associated with better PM), while the measure for updating (*n*-back) was positively correlated with total proportion correct on Virtual Week.

Table 5.4

Correlations between Age, PM Measures and Task Types, and Facets of Executive Functioning

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Age																				
2. CAMP Total	-.13																			
3. CAMP TB	-.17	**.85**																		
4. CAMP EB	-.11	**.74**	**.35**																	
5. MIST Total	**-.32**	**.33**	**.29**	**.28**																
6. MIST TB	**-.27**	**.29**	**.37**	.09	**.72**															
7. MIST EB	**-.21**	**.29**	.19	**.34**	**.81**	**.25**														
8. VW Total	**-.34**	**.25**	.15	**.26**	**.27**	.19	**.24**													
9. VW EB	-.19	.00	-.04	.06	.16	.07	.14	**.80**												
10. VW TB	**-.24**	**.35**	**.27**	**.33**	.19	.21	.17	**.57**	.02											
11. VW TI	-.30	**-.35**	-.20	**-.35**	**-.28**	-.18	-.21	**-.31**	-.17	.18										
12.MEMO Tot	.10	.21	.06	**.31**	.13	.07	.13	**.23**	.14	.13	-.16									
13. MEMO EB	.00	.21	.11	**.26**	.10	.07	.11	**.29**	.14	**.24**	**-.25**	**.93**								
14. MEMO TB	.02	-.01	.02	-.05	.19	**.25**	.09	.09	.01	.07	.02	.17	.20							
15. MEMO TI	.12	-.01	.07	-.13	.11	**.30**	-.04	-.04	-.04	-.08	-.05	.15	.24	-.01						
16. CRT	-.15	.00	.04	-.06	**-.22**	**-.28**	-.10	-.18	-.07	-.07	-.11	-.10	-.08	-.06	-.09					
17. HVLT Tot	-.16	.12	.14	.08	**.20**	.08	.18	**.26**	.06	.06	.08	**.33**	**.28**	-.12	.02	.18				
18. HVLT Ret	-.04	.08	.03	.18	.14	.08	.15	**.22**	.14	.14	**.23**	**.23**	**.36**	.17	.17	-.10	.03			
19. TST	-.08	.01	-.02	.03	.14	.18	.14	.09	.08	.08	.11	-.08	.03	-.07	.04	**.37**	.04	-.07		
20. N-Back	-.13	-.01	.04	-.07	**.26**	.18	.15	**.29**	.16	.16	.18	.08	.18	.06	.11	**.39**	.08	.04	-.02	
21. Go-no go	.02	.00	.02	-.03	-.11	-.07	-.14	-.11	-.13	-.13	**-.27**	**-.21**	**-.30**	.04	-.02	-.06	.06	-.10	-.01	.04

Correlations in bold type $p < .05$.

Note VW TI (Virtual Week time-interval) and MEMO EB (event-based) transformed variables used.

5.6.3 Regression Analyses

A series of two stage hierarchical regression analyses were run with the total and subscale scores for each PM measure. In Stage One measures of age, processing speed, learning, and retention were entered. In Stage Two measures of shifting, updating, and inhibition were entered. Tables 5.5–.5.7 show analyses for the CAMPROMT total, event, and time-based scores, respectively.

Table 5.5

Hierarchical Regression for CAMPROMT Total Score

		Stage One		Stage Two	
	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	-.06	.002	-.04	.002	-.05
CRT	.00	.000	.02	.000	.01
HVLT Tot Rec	.12	.012	.11	.012	.11
HVLT Ret	.08	.004	.06	.005	.07
TST	.01	—	—	.000	-.03
N-back	-.01	—	—	.001	.01
Go-no go	.00	—	—	.001	.03
		$R^2 = .021^{ns}$		R^2 Change = $.001^{ns}$ $R^2 = .022^{ns}$	

ns not significant, p >.05

HVLT Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks; CRT: Choice reaction task, response latencies in milliseconds for all trials.

Table 5.6

Hierarchical Regression with CAMPROMT Event-Based Measure as Dependent Variables

		Stage One		Stage Two	
	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	-.10	.013	-.07	.016	-.08
CRT	-.06	.003	-.06	.010	-.12
HVLT-Tot Rec	.08	.002	.04	.002	.05
HVLT-Ret	.18	.031	.18	.036	.20
TST	.03	—	—	.000	-.01
N-back	-.09	—	—	.020	-.15
Go-no-go	-.03	—	—	.001	.04
		$R^2 = .05^{ns}$		R^2 Change = $.02^{ns}$ $R^2 = .07^{ns}$	

ns not significant, $p > .05$

HVLT-Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT-Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks; CRT: Choice reaction task, response latencies in milliseconds for all trials.

There were no significant predictors for any of the CAMPROMT PM measures. For the CAMPROMT time-based (putatively high demand) score (see Table 5.7), the negative standardized beta value for age, reflecting poorer performance with increasing age, did not diminish from Stage One (-.118) to Stage Two (-.118) when the facets of executive functioning were entered into the model. In contrast, for the CAMPROMT (putatively low demand) event-based score the detrimental influence of age actually increased from Stage One (-.080) to Two (-.094), while the contribution of better processing speed (from -.058 to -.115), learning (from .044

to .047) and retention (from .179 to .198) also increased, indicative of suppressor values or unstable estimates. For all CAMPROMT measures age accounted for a consistent 0-1% of variation in both Stage One and Two; while learning, retention and facets of executive functioning consistently accounted for a meager 0-4% of variation.

Table 5.7

Hierarchical Regression with CAMPROMT Time-Based Measure as Dependent Variables

		Stage One		Stage Two	
	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	-.13	.013	-.12	.013	-.12
CRT	.04	.005	.08	.005	.09
HVLT Tot Rec	.14	.016	.13	.016	.13
HVLT Ret	.03	.000	.01	.000	.01
TST	-.02	—	—	.000	.00
N-back	.04	—	—	.001	.04
Go-no go	.02	—	—	.001	.04
		$R^2 = .04^{ns}$		R^2 Change = $.00^{ns}$ $R^2 = .04^{ns}$	

ns not significant, $p > .05$

HVLT Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks; CRT: Choice reaction task, response latencies in milliseconds for all trials.

The other clinical measure, the MIST, showed a different pattern of results (see Tables 5.8–5.10). The R^2 value (.167) for Stage One of the model predicting outcome on the MIST total raw score was significant, F (4, 75) = 3.77, p = .008. When considering the individual predictors, increasing age was a significant predictor of lower total score on the MIST, accounting for approximately 7% and 6% of variance in Stages One and Two respectively (see Table 5.8). Although the standardized beta value of age was reduced when facets of executive functioning were entered in Stage Two of the model, age continued to be a significant predictor. For the total raw score on MIST, processing speed, learning, retention and facets of executive functioning accounted for only 0-3% of variance.

Table 5.8

Hierarchical Regression with MIST Total-Score as Dependent Variables

		Stage One		Stage Two	
	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	-.32**	.068	-.27*	.058	-.25*
CRT	.22*	.030	-.18	.004	-.08
HVLT-Tot Rec	.20*	.014	.12	.013	.12
HVLT-Ret	.14	.016	.13	.011	.11
TST	.14	—	—	.009	.10
N-back	.26**	—	—	.029	.18
Go-no go	-.11	—	—	.002	-.05
		R^2 = .17**		R^2 Change = .03^{ns} R^2 = .20^{ns}	

ns not significant, p >.05, *p < .05 **p < .01 (two tailed)
HVLT-Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT-Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks; CRT: Choice reaction task, response latencies in milliseconds for all trials.

There were no significant predictors in the model for the event-based (putatively low demand) MIST score, though again the standardized beta value for age was reduced when facets of executive functioning were entered in Stage Two of the model (see Table 5.9).

Table 5.9

Hierarchical Regression with MIST Event-Based Score as Dependent Variables

		Stage One		Stage Two	
	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	-.21*	.027	-.17	.021	-.15
CRT	-.10	.004	-.07	.000	.03
HVLT-Tot Rec	.18	.017	.14	.017	.13
HVLT-Ret	.15	.017	.13	.011	.11
TST	.14	—	—	.018	.15
N-back	.15	—	—	.014	.13
Go-no go	-.14	—	—	.008	-.09
		$R^2 = .09^{ns}$		R^2 Change = $.03^{ns}$ $R^2 = .12^{ns}$	

ns not significant, $p > .05$, $*p < .05$

HVLT-Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT-Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks; CRT: Choice reaction task, response latencies in milliseconds for all trials.

The Stage One model for time-based (putatively high demand) MIST score was significant, F (4, 75) = 3.00, p = .024, accounting for approximately 14% of variance in the outcome measure (see Table 5.10). Age accounted for approximately 5% of variance in outcome in Stage 1 and 2 of the model, however the standardized beta value for age decreased, and was only marginally significant, when facets of executive functioning were entered at Stage Two. Measures of executive functioning accounted for 0-1%, with shifting being relatively more important than updating and inhibition.

Table 5.10

Hierarchical Regression with MIST Time-Based Score as Dependent Variables

		Stage One		Stage Two	
	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	-.27**	.050	-.23*	.045	-.22[t]
CRT	-.28**	.060	-.25*	.025	-.19
HVLT Tot Rec	.08	.000	-.00	.000	-.00
HVLT Ret	.08	.008	.09	.007	.09
TST	.18[t]	—	—	.010	.11
N-back	.18[t]	—	—	.006	.09
Go-no-go	-.07	—	—	.000	-.01
		$R^2 = .14$*		R^2 Change = .01[ns] $R^2 = .15$ [ns]	

ns not significant, *p* >.05, **p* < .05 ***p* < .01 (two tailed), *t* (trend) *p* < .10

HVLT Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks; CRT: Choice reaction task, response latencies in milliseconds for all trials.

The regression model for the Virtual Week total proportion correct was significant at Stage One, *F* (4, 73) = 4.74, *p* = .002, and Stage Two, *F* (7, 70) = 2.94, *p* = .009 (see Table 5.11). Increasing age consistently predicted lower Virtual Week total proportion correct (accounting for approximately 7% of variance at Stage One and Two), though the standardized beta value for age was reduced when facets of executive functioning were entered in Stage Two.

Table 5.11

Hierarchical Regression with Virtual Week Total Proportion Correct as Dependent Variables

		Stage One		Stage Two	
	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	-.34**	.074	-.28*	.066	-.27*
CRT	-.18^t	.017	.13	.003	-.06
HVLT Tot Rec	.26**	.029	.17	.028	.17
HVLT Ret	.22*	.039	.20^t	.031	.18^t
TST	.09	—	—	.003	.06
N-back	.23*	—	—	.019	.15
Go-no-go	-.11	—	—	.001	-.03
		R^2 = .21**		R^2 Change = .02^{ns} R^2 = .23^{ns}	

ns not significant, *p* >.05, **p* < .05 ***p* < .01 (two tailed), *t* (trend) *p* < .10

CRT: Choice reaction task, response latencies in milliseconds for all trials HVLT Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks.

For event-based (putatively low demand) proportion correct in Virtual Week, learning was a significant predictor, accounting for approximately 6% of variance in both Stage One and Two (see Table 5.12). The standardized beta value for age was also identical in Stage One and Two, though only accounting for approximately 2% of variance. There was no significant change in the amount of variance accounted for when entering executive functioning measures in Stage Two.

Table 5.12

Hierarchical Regression with Virtual Week Event-Based Proportion Correct as Dependent Variable

		Stage One		Stage Two	
	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	-.19*	.025	-.13	.022	-.13
CRT	-.16^{t}	.011	-.12	.002	.09
HVLT Tot Rec	.30**	.057	.25*	.057	.25*
HVLT Ret	.15^{t}	.016	.13	.014	.14
TST	.05	—	—	.000	.01
N-back	.16^{t}	—	—	.013	.08
Go-no go	.02	—	—	.000	.09
		$R^2 = .14$*		R^2 Change = .01^{ns} $R^2 = .15^{ns}$	

ns not significant, $p > .05$, *$p < .05$ **$p < .01$ (two tailed), *t* (trend) $p < .10$

CRT: Choice reaction task, response latencies in milliseconds for all trials; HVLT Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks.

For the Virtual Week proportion correct on the time-of-day (putatively intermediate demands) age was marginally significant as a predictor and consistently accounted for approximately 4% of variance in performance (see Table 5.13). The standardized beta value for age diminished in Stage Two when executive functions were included, though executive functions only accounted for an additional 3% of variance in time-of-day performance.

Table 5.13

Hierarchical Regression with Virtual Week Time-of-Day Proportion Correct as Dependent Variable

		Stage One		Stage Two	
	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	-.24*	.046	-.22^{t}	.037	-.20^{t}
CRT	-.07	.002	-.05	.001	.03
HVLT-Tot Rec	.06	.000	.01	.000	.00
HVLT-Ret	.14	.016	.13	.009	.10
TST	.08	—	—	.005	.08
N-back	.16^{t}	—	—	.017	.14
Go-no-go	-.13	—	—	.007	-.09
		$R^2 = .07^{ns}$		R^2 Change = $.03^{ns}$ $R^2 = .10^{ns}$	

ns not significant, p >.05, *p < .05; t (trend) p < .10

CRT: Choice reaction task, response latencies in milliseconds for all trials. HVLT-Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT-Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks.

For the time-interval (putatively high demand) Virtual Week measure, the model for Stage One (including age, processing speed, learning and retention) was significant, F (4, 72) = 2.96, p

= .025, accounting for approximately 14% of variance (see Table 5.14). The inclusion of the executive functioning measures, in Stage Two of the Virtual Week time-interval proportion correct analysis, accounted for an additional 6% of variance. The standardized beta value for age decreased from Stage One (-.276), accounting for approximately 7% of variance, to Stage Two (-.250), accounting for just under 6% of variance. Executive function measures accounted for between 1-4% of variance, with inhibition being a marginally significant predictor.

Table 5.14

Hierarchical Regression with Virtual Week Time-Interval Proportion Correct as Dependent Variable

		Stage One		Stage Two	
	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	-.30**	.073	-.28*	.059	-.25*
CRT	-.11	.007	-.08	.000	.02
HVLT Tot Rec	.08	.000	.00	.000	-.00
HVLT Ret	.23*	.047	.22^{t}	.026	.17
TST	.11	—	—	.009	.10
N-back	.18^{t}	—	—	.019	.15
Go-no go	-.27**	—	—	.039	-.20^{t}
		R^2 = .14*		R^2 Change = .06^{ns} R^2 = .20^{ns}	

ns not significant, *p* >.05, **p* < .05 ***p* < .01 (two tailed), *t* (trend) *p* < .10

CRT: Choice reaction task, response latencies in milliseconds for all trials. HVLT-Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT-Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks.

The regression model for the total proportion correct on MEMO was significant at Stage One, *F* (4, 58) = 3.62, $p = .011$, and accounted for 20% of variance (see Table 5.15). However, the addition of the executive function variables did not significantly improve the model, accounting for only an additional 2.7% in overall proportion correct. The only variables that predicted PM performance overall on the MEMO paradigm was learning and retention. In particular, learning was a robust predictor that explained between 10 and 11% of variance in overall performance. Inhibition and age were both marginally significant ($p < .10$) predictors and the relevance of age as a predictor slightly decreased in Stage Two.

Table 5.15

Hierarchical Regression with MEMO Total Proportion Correct as Dependent Variables

		Stage One		Stage Two	
	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	.10	.042	.22^{t}	.040	.21
CRT	-.10	.024	-.16	.012	-.12
HVLT-Tot Rec	.33**	.104	.33**	.113	.35**
HVLT-Ret	.23*	.055	.24*	.027	.18
TST	-.08	—	—	.001	.04
N-back	.08	—	—	.002	.03
Go-no go	-.21*	—	—	.026	-.18^{t}
		$R^2 = .20$*		R^2 Change = .03^{ns} $R^2 = .23^{ns}$	

ns not significant, $p > .05$, *$p < .05$ **$p < .01$ (two tailed), *t* (trend) $p < .10$

CRT: Choice reaction task, response latencies in milliseconds for all trials; HVLT-Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT-Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks.

Table 5.16

Hierarchical Regression with MEMO Event-Based Proportion Correct as Dependent Variables

		Stage One		Stage Two	
Variable	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	.00	.011	0.11	.012	0.11
CRT	-.08	.022	-.15	.002	0.08
HVLT Tot Rec	.28**	.062	0.25*	.074	0.30*
HVLT Ret	.36**	.127	0.37*	.065	0.27*
TST	.03	—	—	.010	0.07
N-back	.18^{t}	—	—	.022	0.16
Go-no-go	-.30**	—	—	.051	-0.29*
		R^2 = .22**		R^2 Change = .07^{ns} R^2 = .29**	

ns not significant, *p* >.05, **p* < .05 ***p* < .01 (two tailed), *t* (trend) *p* < .10

CRT: Choice reaction task, response latencies in milliseconds for all trials; HVLT-Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT-Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks.

For the MEMO event-based (putatively low demand) proportion correct, learning, retention, and inhibition were all significant predictors (see Table 5.16). The inclusion of executive functions in the model accounted for an additional 7.2% of variance in event-based proportion correct scores. Age was not a significant predictor in Stage One and accounted for only just over 1% of variance in performance in event-based MEMO proportion correct. This did not change in Stage Two. Inhibition was a significant predictor in the expected direction and accounted for 5% of variance in outcome. Shifting and Updating accounted for only 1 and 2% of variance respectively. Most notably in Stage Two, learning and retention accounted for over 7 and 6% of variance respectively.

Table 5.17

Hierarchical Regression with MEMO Time-of-day Proportion Correct as Dependent Variables

		Stage One		Stage Two	
Variable	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	.02	.001	.03	.001	.04
CRT	-.06	.010	-.10	.015	-.14
HVLT-Tot Rec	-.12	.018	-.14	.025	-.16
HVLT-Ret	$.17^{t}$	.041	.21	.050	$.25^{t}$
TST	-.07	—	—	.010	-.10
N-back	.06	—	—	.001	.03
Go-no go	.04	—	—	.017	.14
		$R^2 = .06^{ns}$		R^2 Change = $.02^{ns}$ $R^2 = .08^{ns}$	

ns not significant, $p > .05$, *t* (trend) $p < .10$

CRT: Choice reaction task, response latencies in milliseconds for all trials; HVLT-Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT-Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks.

Models for the time-of-day MEMO proportion correct were non-significant (see Table 5.17), with both age and executive functioning measures apparently irrelevant (0-1% of variance). Retention was marginally significant as a predictor in Stage Two and accounted for 5% of variance.

For the time-interval (putatively high demand) MEMO proportion correct score analyses (see Table 5.18) there was no evidence for meaningful age-effects on this task type (which in this study permitted participants' use of external aids), or for executive function mediating the age-effects on this task type.

Table 5.18

Hierarchical Regression with MEMO Time-Interval Proportion Correct as Dependent Variables

		Stage One		Stage Two	
Variable	Raw Correlation	Squared Semi-Partial Correlation	Standardised Regression Coefficient	Squared Semi-Partial Correlation	Standardised Regression Coefficient
Age	.12	.026	.17	.030	.18
CRT	-.09	.022	-.16	.012	-.13
HVLT Tot Rec	.02	.000	.02	.000	.01
HVLT-Ret	.17[t]	.043	.21	.044	.23
TST	.04	—	—	.001	.09
N-back	.11	—	—	.007	.04
Go-no go	-.02	—	—	.003	.06
		$R^2 = .07^{ns}$		R^2 Change = $.01^{ns}$ $R^2 = .08^{ns}$	

ns not significant, $p > .05$

CRT: Choice reaction task, response latencies in milliseconds for all trials; HVLT-Tot Rec: Highest number of items recalled on any of the first three immediate recall trials; HVLT-Ret: Hopkins Verbal Learning Test Revised retention percentage score TST: Task switch task; N-back score is for reaction time for hits; Go-no go score is total number of commission errors over all blocks.

5.6.4 Mediation Analyses

There was a significant indirect effect for the MEMO total proportion correct score through learning (i.e., highest recall of words on first three immediate recall trials of HVLT-R), (-0.057) (0.030) = -0.002 (bootstrap Standard Error [*SE*] = 0.001; bootstrap Confidence Intervals [CI] -0.005, -0.000). The completely standardized indirect effect was -.072 (*SE* = 0.042; CI -0.181, -0.003). A second significant indirect effect was also found for the MEMO event-based proportion score with learning, (-0.057) (0.027) =0.003 (bootstrap SE =0.001; bootstrap CIs -0.004, 0.000). The completely standardised indirect effect was -0.067 (SE = 0.044; CI -0.185, -0.002). None of the other indirect effects were significant. A serial mediation model was also assessed using Mplus Version 8 (Muthen & Muthen, 2017) with processing speed as first mediator and facets of executive functioning as second mediators, however, none of the path analyses showed any significant indirect effects.

5.7 Discussion

Understanding the mechanisms that lead to normal age-related decline on diverse PM tasks is important for developing targeted support for older adults in the community. The present study continues and extends previous work on facets of executive functioning, and other potential mediators of the relationship between age and PM performance, by investigating the applicability of the hypothesised mediational role of these facets to a wide range of PM tasks (in the clinic, the laboratory, and naturalistic-settings). The present study examined how three common executive functioning measures (which are recognised as largely mapping onto the three main facets of shifting, updating, and inhibition), predict PM performance. It was hypothesised, based on the multiprocess framework, the four phase process model, and previous empirical research on age-effects on PM in the laboratory, that facets of executive functioning would primarily mediate high demand PM tasks. The main findings were that age was the best predictor of overall PM performance on one clinical measure (MIST) and a laboratory measure (Virtual Week), even after measures of retrospective memory (learning and retention), processing speed, and facets of executive functioning were controlled for. The mediation analyses revealed that increasing age was related to small decrements in verbal learning, and that verbal learning in turn was related to better PM performance for the total and event-based proportion correct on the MEMO. The only facet of executive functioning that was a significant predictor was inhibition, which predicted the MEMO (putatively low demand) event-based proportion correct. Inhibition was also marginally significant for predicting Virtual Week (putatively high demand) time-interval proportion correct scores. There was no evidence that facets of executive function mediated age-effects more on the high compared to low demand PM tasks.

Key to formulating this investigation was the main tenant of the multiprocess model (McDaniel & Einstein, 2000) and the four phase process model (Kliegel et al., 2002), which proposes that different PM tasks are more or less demanding of facets of executive functioning. In particular, the multiprocess model holds that event-based PM tasks that are salient or related to the processing involved in the ongoing task are more likely to be mediated by relatively spontaneous processes, such as occur in recognition memory paradigms. In contrast the four phase process model suggests that executive functioning is involved in all phases of a PM task; except the maintenance of the intention during the delay, where retrospective memory processes (retention) are expected to play a critical role. It was noted that both of these models are relatively mute about which particular facets are involved in demanding event and time-based

PM tasks. Thus, much previous research on facets of executive functioning mediating age-effects on PM has been exploratory and somewhat speculative, as has the present study's approach.

In comparison to recent studies by Zuber et al. (2016) and Schnitzspahn et al. (2013), the present findings do not support shifting and updating mediating age-effects on time-based PM tasks. Intuitively, it seems that an intention is more likely to be successfully executed when refreshed intermittently with updating of outstanding tasks to be performed, particularly on time-based tasks where monitoring appears to be especially important. The present study did not find convincing support for this conceptualisation of the role of updating in the association between age and PM, though updating ability was positively correlated with overall performance on the total Virtual Week proportion correct score and the total MIST raw score. When entered in the second stage of the regression models for each PM outcome measure, updating accounted for only 1–3% of unique variability in PM. This is a stark contrast with the finding of Rose et al. (2010) of a high correlation (r =.50) between older adults' updating ability and the time-interval proportion correct on the Virtual Week PM task. Somewhat surprisingly, in the present study, age did not appear to correlate strongly with the measures of executive functioning used. This lack of correlation between age and executive functioning represents a possible limitation of the current findings and casts some doubt on whether the other correlations between age and PM can be interpreted with confidence. However, it should also be noted that while executive functions are often expected to correlate with age in the literature, they sometimes do not (Henry & Phillips, 2006; Phillips et al., 2011; Phillips & Henry, 2010).

In formulating their hypotheses for facets of executive functioning potentially mediating association between age and PM performance on low demand (focal) PM tasks, Zuber et al. (2016) argue that inhibition appears to be especially related to focal event-based PM tasks in which it is assumed that there is less need to break set from an ongoing task to monitor for a cue (and as a result a greater need for inhibiting a prepotent ongoing response when the anticipated cue does appear). In the present study inhibition, like the other facets of executive functioning, was mostly unrelated to performance on the PM outcome measures. Though, consistent with the argument of inhibition being necessary to break set with a habitual or relatively automatic ongoing task, inhibition did show a significant negative correlation with the MEMO event-based proportion correct score (i.e., higher inhibition ability was related to better PM performance on this measure). Most fitting for this interpretation, the MEMO had as its ongoing task the habitual and daily routine activities of the older adults who participated in this naturalistic-setting PM task.

An important difference between the present study, and much of the previous research on executive functioning ability mediating age-effects on PM performance, is that the majority of those previous studies included younger adults (e.g., Zuber et al., 2016) or a young adult group (e.g., Cherry & Lecompte, 1999). It might be natural to assume from this that the present study had a restricted age range that prevented the mediating role of facets of executive functioning appearing in age-effects on PM. However, in response to this it can be noted that most cognitive decline occurs later in life (Park & Reuter-Lorenz, 2009) and so it could equally be argued that the relevance of executive functioning ability should be as apparent, or more so, when comparing young-old and old-old adults. Instead, the more likely explanation for the contrast in findings between the current and previous studies is the difference in PM measures used. Specifically, the classical laboratory paradigm used in previous studies is characterised by relatively abstract ongoing tasks and low environmental supports, whereas the PM tasks used in the present study are more akin to typical daily life PM tasks (cf. Kliegel et al., 2007; Kliegel, Rendell, et al., 2008).

For example, the low demand PM task in Zuber et al. (2016) was to press the space bar when the letter “A” or “D” appeared within a *n*-back ongoing task, and the high demand task was to respond when the frame for the stimuli changed a particular colour. The present study had more familiar low demand PM tasks, such as the event-based task in the MIST: “when I hand you an envelope, self-address it”, within an ongoing task (word search in the case of the MIST) which many older adults seemed to find familiar. Only some of the high demand tasks in the present study were akin to the classic PM paradigm. Such as the time-interval task in Virtual Week, which involved clicking a button marked “Perform Task” and choosing from a list of activities “test lung capacity” at two and four minutes on a stop clock. Thus careful consideration of the PM and ongoing task differences both *between* and *within* the classic and more ecologically or “familiar” paradigms seems warranted to further illuminate when age-effects are likely to manifest (cf. Kliegel et al., 2007).

Given the relative paucity of research on different facets of executive functioning mediating age-effects on a range of PM measures, the current research was largely exploratory in its aims. The results were mostly inconsistent with the general hypothesis that facets of executive functioning mediate age-effects on putatively high (and low) demand PM tasks across clinical, a simulated naturalistic laboratory paradigm (Virtual Week), and a naturalistic-setting PM paradigm with a range of task types (MEMO). The indirect age-effect on PM of processing speed alone, and mediating performance on facets of executive functioning, was also not supported. It might be speculated that one reason for processing speed not being relevant for explaining age-effects on PM in the current study is that many of the PM measures used did not have the same time pressure or constraints as the classic PM paradigm, used by other studies that have investigated the age-effects being mediated by facets of executive functioning. Although this is the case for some of the PM tasks used here, it is not the case for others. For example, Virtual Week included a time interval measure that was similar to other time-based tasks used in the laboratory, as did the MEMO, albeit with longer delays. However, in the case of event-based tasks time demands were less constrained making facets of executive functioning less relevant. The cue for the event-based tasks in the clinical measures, Virtual Week, and MEMO, was presented for up to a minute in the case of the MIST and potentially up to an hour or more in the case of the event-cues in MEMO (e.g., taking a photo of lunch). It is noteworthy in this respect that learning and retention were better predictors of PM performance than facets of executive functioning.

Some support for executive functioning mediating the relationship between age and at least some PM tasks was evident in the regression modelling involving the MIST. Specifically, analysis showed that age was a significant predictor of overall performance on the MIST, independent of processing speed, learning and retention. However, this age-effect was diminished when measures of facets of executive functioning were entered—though the latter measures themselves did not significantly predict PM performance and the amount of variance explained by age only decreased by 1–3% in this second stage. Some evidence for processing speed mediating age-effects on PM via executive functioning was also shown in the MIST time-based score, where processing speed was a significant predictor in the first stage but not second stage of regression model (beta for each stage for processing speed was -.250 and -.185). There was also evidence of a suppression effect for age, with the beta values for age increasing slightly in some of the regression models (e.g., MEMO time-based tasks, and CAMPROPMT total and event-based scores).

The present findings of learning mediating age-effects on PM for the MEMO total and event-based proportion correct score is partially consistent with the previous findings of Gonneaud et al. (2011), who found performance on a measure of retrospective memory and learning fully mediated older adults performance on event-based PM tasks. However, the present study has extended these previous findings by showing that a similar relationship appears to hold for event-based tasks in a naturalistic-setting. Although only accounting for a small amount of variance in PM performance, the present findings suggest that learning and retention may be as important as facets of executive functioning in accounting for older adults PM performance on more naturalistic PM measures. Perhaps this is not surprising given that the literature on older adults' PM performance in naturalistic-settings consistently shows superior performance compared to younger adults (Henry et al., 2004). There is therefore a range of possible mediators for age-effects on PM performance in naturalistic-settings apart from facets of executive functioning (Azzopardi et al., 2015) that requires further investigation. For example, further investigation on how older adults compensate for age related declines in various cognitive abilities (Gilbert, 2015; Park & Reuter-Lorenz, 2009) may be one fruitful avenue for further research. Future research could build on the present study's approach by continuing to explore the range of PM task types across paradigms and settings and how they relate to multiple potential mediators of age-effects on PM. It should also be noted that the current study, while staying close to the Miyake et al. (2000) model of core executive functions, did not include other constructs that may also be included as executive functions, such as planning (e.g., Kliegel, Martin, McDaniel, & Phillips, 2007) and meta-cognition (e.g., Hertzog & Dunlosky, 2011). Thus future studies may also extend the current research by exploring these other executive functioning constructs in relation to the range of PM measures used in the current study.

In sum, the present study found that age was a better predictor than processing speed, aspects of retrospective memory, and facets of executive functioning for the total score performance on the MIST and Virtual Week, and that verbal learning mediated older adults' performance on the MEMO total and event-based proportion correct score. Facets of executive functioning were consistently unrelated to performance on the diverse measures of PM used in the present study. This finding suggests that the theoretical assumption and previous empirical findings that facets of executive functioning mediate older adults declines on high (and possibly low) demand PM task types needs to be considered in light of possible boundary conditions, such as the PM paradigm used to investigate this relationship. Further work on how age-effects manifest in performance on a diverse range of PM task types is clearly warranted.

Chapter 6: General Discussion

The previous chapters have brought out the multifaceted nature of PM and its complex relationship to the ageing process, particularly in the later phases of the lifespan. This final chapter presents a review, synthesis, and critical discussion of the findings reported in this thesis. First, to provide appropriate structural cohesion with the introductory chapter, section 6.1 will review the aims and rationale for this research project. Section 6.2 gives a concise review of the ageing and the PM literature to provide context for the following sections. In section 6.3 the current research findings are critically considered within the context of previous research findings. The implications of the present research findings are discussed in relation to the current theories of ageing and PM in section 6.4. Section 6.4.4 considers how methodological considerations are critical in investigating ageing and PM. Such considerations include the necessity to develop conceptually parallel tasks across settings; the importance of considering the concept of time in PM tasks; and the variability of measures purporting to measure executive functioning ability and what this means for hypothesis testing. The following section 6.5 considers the strengths and limitations of the present thesis, which is then followed by a section giving a concise presentation of the unique contributions of the present thesis to the ageing and PM field (section 6.6). In the final two sections (6.7 and 6.8) future directions for research within the ageing and PM field are presented and discussed, based on the critical consideration of the present research project and findings, and a conclusion is presented.

6.1 Recapitulation of Research Questions and Aims

The main research questions motivating the current research project were how cognitive ageing interacts with PM task characteristics, such as level of environmental support; and whether individual differences in controlled attentional processes within older adults mediates PM task performance when environmental support is minimal. More specifically, the current research aimed to investigate whether an absence of conceptually parallel PM task types, differing in levels of environmental support, may account for the age-PM paradox (i.e., the pattern of young outperforming older adults in laboratory-settings, and vice versa in naturalistic-settings; and young-old outperforming old-old in laboratory-settings but showing no differences in naturalistic-settings). Key to this investigation was distinguishing different types of time-based tasks that, it was argued, are often inappropriately compared across settings: time-of-day (appointment-like tasks; e.g., doctor appointment at 11am) and time-interval tasks (elapsed time intervals as cues to perform a PM task; e.g., taking a cake out of the oven in 25 minutes). The second specific aim was to investigate individual differences in cognitive functioning, within a healthy older adult cohort, as a possible explanation for why memory declines for PM may be relatively small for some older adults but larger for others. The third specific aim was to consider how general changes and individual differences in cognitive ageing interact with PM task demands, such that PM performance may be relatively unaffected when cognitive demands are low, but more affected when cognitive demands are high. The final specific aim was to begin to elaborate on and make clearer a distinction within time-based tasks that may have the same utility as the focal versus non-focal event-based cue distinction within many laboratory studies.

6.2 Cognitive Ageing and Prospective Memory

The Chapter 2 literature review critically considered four of the current main theories in the ageing-PM field.[19] The self-initiated operations model of PM (Craik, 1986) begins from a broad consideration of age-effects on laboratory memory tasks in general. In particular it draws attention to the interaction between cognitive processes and the environment, which enable older adults to perform well on some memory tasks (e.g., recognition tests) when environmental support is high, but leads to poorer performance when environmental support is low (e.g., free recall tests). Extending this logic to PM, the self-initiated operations model noted that there is no explicit prompt to recall an intention that is to be executed in PM tasks (almost by definition); and that therefore all PM tasks should show more age-effects than even low support retrospective memory tasks, like free recall where the researcher does prompt the participant. The multiprocess framework subsumed the central tenant of the self-initiated operations model, but also made room for relatively spontaneous retrieval processes in PM tasks (McDaniel & Einstein, 2000). It conceptualised PM as involving multiple retrieval processes which are recruited as a function of characteristics of the PM task (particularly type of cue, such as event or time-based) and individual differences (including age). This extension provided an explanation for apparently anomalous findings in the ageing and PM literature; where older adults showed no PM deficits relative to young adults on some event-based tasks in the laboratory (this model does

[19] As noted in the chapter 2 footnote on p.48, Delay Theory (Heathcote et al., 2015; Strickland et al., 2018) may prove of some interest to researchers working in the ageing and PM field in the future. In particular, the decision process that it mathematically models may show some variation with the ageing process, such as higher thresholds of evidence accumulation prior to making decisions (responses) in event-based PM tasks with age—leading to similar accuracy performance but longer latencies relative to younger adults in the classic experimental PM paradigm. The emergence of such different strategic biases in decision making processes may show some kinship with changes in context representation with ageing (Braver & Barch, 2002; Braver et al., 2001; Braver et al., 2003; Braver et al., 2005). However, the fruitfulness of Delay Theory, in terms of developing research questions and empirical knowledge, has yet to be shown in the ageing and PM field and thus has not figured as more than a footnote in this thesis.

not explicitly consider time-based tasks or older adults' superior performance compared to young adults on time-of-day tasks in naturalistic-settings). In turn, the multiprocess framework was further developed into the dynamic multiprocess framework, in which a dynamic interplay of self-initiated operations and relatively spontaneous retrieval processes occurs in PM tasks; especially those in naturalistic-settings with longer delays between forming an intention and executing it (Shelton & Scullin, 2017). This model highlighted the importance of considering contextual cues (e.g., being in the office) as leading to spontaneous retrieval of previously formed intentions (e.g., passing on a message to a colleague), which in turn leads to controlled attentional processes, such as monitoring for cues to execute the PM task (e.g., seeing the colleague to whom the message is to be passed on coming out of a meeting). Finally, the preparatory attention and memory (PAM) processes model (Smith, 2008) proposed that cognitive resources are always taxed when maintaining a delayed intention and monitoring for the cue to execute it. This might not always be obvious, as often a trade off on performance on the ongoing task (e.g. increased latencies in responding) is made to maintain adequate PM performance (Smith et al., 2012). Each of the preceding models extends the range of age associated PM phenomena that can be accounted for, as well as providing clear hypotheses that can be tested. However, the theoretical distinction between time-based tasks (time-of-day and time-interval), explored in the present research project, has not figured in any of these theories, and this lacuna is significant, given no attempt so far has satisfactorily accounted for the age-PM paradox (Henry et al., 2004; Rendell & Thomson, 1999).

Thus, although there is empirical support for each of the models of PM discussed in this thesis, it is proposed here that all are challenged by the robust pattern of discrepancies between young and older, and young-old and old-old, adults' performance across research settings (i.e., both aspects of the age-PM paradox). A natural corollary to trying to make sense of the age-PM paradox is to consider the methodology (Uttl, 2008) and types of PM tasks used across both settings (Henry et al., 2004). The first empirical study in this thesis was motivated by the aim to control for potential methodological drivers of the age-PM paradox. In particular, PM task types were designed to be conceptually equivalent across settings, and a key distinction between time-of-day (relatively high environmental support) and time-interval (relatively low environmental support) time-based PM tasks was also made. It was proposed that this latter distinction may be of a similar significance to that within event-based PM tasks between focal (low monitoring demand) and non-focal (high monitoring demand) cues.

Previously, Rendell and Craik (2000) found that a time-interval task in Virtual Week and Actual Week (perhaps the task type that is most practically amenable to parallel implementation in both laboratory and naturalistic-settings) led to similar performance by older, but not younger, adults across settings. In Virtual Week, the proportion correct on the time interval task for young, young-old, and old-old were .72, .47, and .34, respectively; while in Actual Week “performance was U-shaped with age” (Rendell & Craik, 2000, p. 57), with proportions correct for young, young-old, and old-old: .24, .46, and .26, respectively. In contrast for both event and time-of-day tasks, primarily considered in terms of the dimension of regularity, there was a marked contrast across settings. In Virtual Week (laboratory-setting) the young, young-old, and old-old proportion correct for irregular and regular event and time-based tasks ranged from .70–.93, .41–.89, and .22–.82, respectively; while in a naturalistic-setting: .51–.68, .74–.81, and .69–.85, respectively, were reported. However, potential limitations of the first attempt by Rendell and Craik (2000) to develop parallel tasks across settings, particularly for event-based and time-interval tasks were raised in Chapter 4. Specifically, time markings on the squares of the board in the Virtual Week version used made the time-of-day tasks practically indistinguishable from event-based tasks. Furthermore, the difficulty in verifying performance on event-based tasks in Actual Week meant that the researchers were overly reliant on the conscientiousness of participants to adhere to instructions for the PM tasks (e.g., to report, using a time-stamped voice recorder, the first time they opened the fridge). The version used of Virtual Week in Chapter 4 overcame the first limitation, while the MEMO paradigm, with its use of date and time-stamped photos, went a considerable way to overcoming the second limitation.

One common thread of the four models of PM reviewed above, and made particularly explicit in the process model discussed in Chapter 5, is the plausible assumption that age-effects on PM (particularly after the age of 60) should be moderated by, or even fully mediated by, declines in controlled processing or facets of executive functioning. This strong assumption has its origins in one of the earliest of the contemporary models of PM, the self-initiated operations model (Craik, 1986). However, the multiprocess framework proposed this was not the case for low cognitive demand tasks, as evidenced by studies where older adults showed little or no PM deficits relative to young adults (Henry et al., 2004). The four phase process model recognizes the insight of the multiprocess framework in regards to some cues (e.g., distinctive or focal) eliciting relatively spontaneous retrieval of a previously formed intention, though this applies only to when the right kind of cue is encountered. However, as distinctively noted by the process model, there are other phases to a PM task, and at least two of these appear to require executive functioning ability: at the encoding of the intention phase and at the execution phase).

Recently, there has been some evidence that it may be only some facets of executive functioning (or that these facets make differential contributions depending on characteristics of the PM task) which are implicated in successfully performing PM tasks. In particular, there is some support for inhibition and shifting ability (which tend to decline with age) being predictors of age-effects on PM (Schnitzspahn et al., 2013; Zuber et al., 2016). Another facet of executive functioning ability, updating, does not appear to be reliably associated with age-effects on PM (Rose et al., 2010; Schnitzspahn et al., 2013), but is still an important candidate for further investigation as at least one study has found that it accounts for a substantial amount of variance in PM performance for older adults on the Virtual Week paradigm (Rose et al., 2010). Another motivation for testing the relevance of executive functioning ability variables in the current research project is that previous studies of their relation to age-effects on PM have tended to be limited to versions of the classic laboratory paradigm of PM (Joly-Burra et al., 2018; Schnitzspahn et al., 2013; Zuber et al., 2016). Thus, there was a clear need to test the executive functioning hypothesis of age-effects on PM by using a range of PM measures across different settings (laboratory, clinical, and naturalistic).

6.3 Current Research Findings

6.3.1 Examining the age-PM paradox with parallel PM task types

The aim of the two studies presented in Chapter 4 was to attempt to account for the *two* aspects of the age-PM paradox—young outperforming older adults in the laboratory but vice versa in naturalistic-settings; and young-old outperforming old-old in laboratory studies but showing no difference in naturalistic-setting studies (Henry et al., 2004; Rendell & Craik, 2000; Rendell & Thomson, 1999). It was hypothesised that the age-PM paradox could be explained by previous studies using discrepant PM task types across settings, particularly in regards to time-based tasks. It was argued in Chapter 4 that the two time-based tasks that have not been sufficiently separated in theory and practice were: time-of-day tasks, involving a constellation of social and activity related time-based cues found in abundance in daily life (Rabbitt, 1996); and time-interval tasks, involving monitoring relatively arbitrary durations of elapsed time. Time-interval tasks it should be noted also frequently occur in daily life (e.g., moving your car in 15 minutes to avoid a fine) but are primarily studied in the laboratory, and, critical to the current research project's argument for understanding the age-PM paradox, are frequently *conflated with all time-based tasks*. It was also noted in Chapter 4 that event-based tasks, on which older adults perform at similar levels to younger adults when the cue is focal or salient, have also been relatively neglected in naturalistic studies, and that for the sake of completeness these tasks should also be made conceptually parallel across setting if possible (Bailey et al., 2010; Niedźwieńska & Barzykowski, 2012). The use of the MEMO paradigm (Randall, 2016) allowed the testing of conceptually parallel tasks to those used in the laboratory paradigm Virtual Week, including both time-based and event-based tasks.

The first experiment presented in Chapter 4 showed the general pattern of young outperforming older adults in the laboratory. However, using the MEMO paradigm in a naturalistic-setting, older adults outperformed younger adults *only on the time-of-day PM task*, consistent with the current research project's diagnosis of the primary cause of the age-PM paradox being due to conflating the two types of time-based tasks. Performance on the event-based task in the naturalistic-setting, using the MEMO paradigm, was similar for both young and older adults (both showed high performance consistent with the general, putatively low cognitive demand for event-based tasks). Young and older adults also showed similar performance on the MEMO time-interval task (both showed low performance consistent with the putatively high cognitive demand for *this* type of time-based task). Most notably, older adults showed a similar level of performance on the time-interval task on MEMO to the time-interval (and time-of-day) tasks used in Virtual Week, supporting the conceptually parallel nature of the time-interval task.

There was also a marked contrast for older adults between time-of-day and time-interval tasks on the MEMO, supporting the distinction between these two time-based tasks.

The results of the second experiment presented in Chapter 4 showed young-old adults outperforming old-old adults in the laboratory. As with Experiment 1 there was no significant difference in the laboratory between the two time-based tasks, which may therefore appear to be indistinguishable in terms of cognitive demand as inferred from the substantially lower level of performance on these tasks compared to the event-based tasks. However, the performance of both older groups on the Virtual Week time-of-day tasks, while lower than proportion correct on event-based tasks (.20 versus .30, respectively), was higher than proportion correct on the time-interval tasks (.20 versus .14, respectively), suggestive of the theoretically “intermediate” cognitive demands of time-of-day tasks. For the naturalistic-setting PM tasks in this second study, the MEMO was made more challenging than the version used in Experiment 1 (i.e., all task types were combined in the same three day period), though this manipulation did not reduce PM performance, which was actually at a higher level than that shown in Experiment 1 (likely due to external aids being permitted, including setting alarms or making notes). There was no difference between young-old and old-old on any of the tasks in the naturalistic-setting. However, both groups of older adults performed better on event-based tasks compared to the time-of-day tasks and time-interval tasks (overall proportions correct, .85, .70, and .71, respectively), almost mirroring the pattern of results in the laboratory, albeit at higher levels of performance across all tasks. The second experiment was completely consistent with the second aspect of the age-PM paradox. It also demonstrates that healthy older adults are capable of finding ways to adapt or compensate for normal age-related cognitive decline in naturalistic-settings.

Both Experiments 1 and 2, presented in Chapter 4, captured some contextual information about participants who performed the scheduled (time-of-day) and random (time-interval) quizzes. The context captured was the participant's location, and the activity the participant had just been engaged in, when they switched to performing the PM task. In Experiment 1 it was found that older adults tended to be more often at home and switching from chores when responding to time-of-day quizzes compared with younger adults, who more often reported being at work. The home location of older adults who successfully responded to the time-of-day quiz provides equivocal evidence for the common assumption that older adults are less busy than younger adults. For example, older adults reported more often being engaged in chores than younger adults, but without knowing the nature of these chores it is hard to draw any firm conclusions about how absorbing or demanding these were. A similar pattern emerged for the time-interval task. In Study 2, there were relatively few differences between young-old and old-old in terms of the location or activities they were engaged in immediately prior to switching to perform the PM tasks. Overall, the results show that healthy older adults' vulnerability to PM lapses (forgetfulness) in naturalistic-settings is relatively confined to time-interval tasks in which external aids or strategies are not available.

6.3.2 The role of facets of executive functioning in diverse low and high demand PM tasks

The models of age-effects on PM considered in this research project, and a considerable body of empirical evidence using primarily the classic Einstein and McDaniel (1990) paradigm, make a strong case for executive functioning mediating age-effects on high demand PM tasks (e.g., non-focal event-based tasks). However, it is not clear whether individual differences in executive functioning mediates age-effects on other PM measures (it certainly does not *seem* to do so in naturalistic-settings, as illustrated by the robust age-PM paradox). The dominant paradigm in the age-PM literature is the Einstein and McDaniel (1990) paradigm. This paradigm is characterised by a number of features, but most strikingly by the typically abstract and unfamiliar ongoing and PM task (e.g., lexical decision making with a PM task of a different key press for target words or categories) and the small number of PM targets (usually no more than 3–4). This approach has the virtue of a high level of experimenter control. For example, using the classic paradigm it is relatively easy to manipulate high and low demand PM tasks by changing the relationship of a PM target cue to the ongoing tasks (Breneiser & Mcdaniel, 2006), as in the case of investigating age-effect on focal versus non-focal PM event-based tasks (Kliegel, Jäger, et al., 2008). However, the small number of PM tasks in the classic PM paradigm poses a concern for reliability (Kelemen, Weinberg, Alford, Mulvey, & Kaeochinda, 2006), particularly when investigating individual differences (Rose et al., 2010) rather than group differences (the latter appears to be more appropriate for investigation with the classic paradigm). A less obvious limitation of the classic PM paradigm is that it is liable to obscure the need to investigate the role of controlled attentional processes (executive functions) in more ecologically valid PM measures.

In Chapter 5 a very novel investigation into the role of facets of executive functioning ability, and other potentially related cognitive processes (i.e., processing speed, learning, and retention), in mediating age-effects on four key PM measures (laboratory, two clinical, and a naturalistic-setting) was carried out for the first time. The unique contribution of this study is particularly noteworthy, given that no studies to date have included both the MIST and CAMPROMT (two widely used clinical measures) in a single study, along with a laboratory PM paradigm with high ecological validity (Virtual Week), and a comprehensive naturalistic-setting

PM paradigm (MEMO). The sample consisted only of healthy older adults from 60–87 years of age and was not subdivided into young-old and old-old, as the focus was on individual differences within a cohort of older adults. PM task types were conceptualised as on a continuum from low to high cognitive demand primarily based on the degree of environmental support inherent in the task, and thus the presumed degree of controlled attentional processing (which tends to decline with age) required as a corollary. Generally low demand PM tasks included event-based (e.g., self-addressing an envelope when it was presented in the MIST, or taking a photo of lunch at that meal time in the MEMO) and time-of-day tasks (e.g., use asthma inhaler at 11am and 9pm, in the Virtual Week game, or opening the MEMO app for scheduled quizzes at 10am and 3.30pm). High demand tasks were those where environmental support was low and controlled attentional processing demands were high. The most obvious high demand PM tasks meeting this criteria were the time-interval tasks. Examples from the measures used include: check lung capacity at 2 and 4 minutes on stop clock in Virtual Week; write down number of medications being taken in 15 minutes time on the MIST; changing tasks in 9 minutes time on the CAMPROMT; and completing a quiz in 10 minutes time on the MEMO. However, the conditions making for high and low demand PM tasks were also noted in Chapter 5 as involving more complex considerations. For example, the ongoing tasks within which a PM task is embedded can vary in how demanding or absorbing of attentional resources they are (Marsh & Hicks, 1998). It was proposed in both Chapter 4 and 5, based on the previous literature on the age-PM paradox, that the ongoing task in laboratory-settings may generally be more demanding (e.g., less familiar or more experimenter-given tasks given in a short space of time) than those presumably routine or habitual ongoing tasks older adults may be engaged in in naturalistic-settings.

For the CAMPROMT (Wilson et al., 2005), a widespread clinical measure of PM, neither facets of executive functioning, processing speed, learning or retention mediated the association of age with PM performance on low (event-based) or high demand (time-based) PM tasks. No mediation via facets of executive functioning, processing speed, learning or retention, was also found for low and high demand PM tasks on the MIST (Raskin et al., 2010), the second clinical measure used. However, age and processing speed was found to be associated with overall performance on the MIST, with increasing age and slower processing associated with decrements in PM performance. Similarly, for Virtual Week (Rendell & Craik, 2000; Rendell & Henry, 2009) age showed a robust association with overall performance. On the event-based Virtual Week tasks (low demand) learning showed some association with performance. Learning and retention was also associated with event-based (low demand) and overall performance on the MEMO. Inhibition was also associated with better event-based (low demand) PM performance on the MEMO. This isolated finding, of inhibition being related to performance on the MEMO event-based tasks, fits with the hypothesis of Zuber et al. (2016) that, for low demand event-based (non-focal) PM tasks, where the ongoing task has a pre-potent quality (e.g., habitual or routine, as might be expected in daily life), being able to inhibit habitual responding on predictable ongoing tasks ought to be particularly important. Finally, the inclusion of other cognitive variables hypothesised to be related to PM performance in general (i.e., processing speed and retrospective memory processes), revealed a small and isolated indirect effect of learning mediating age-effects on PM performance on the MEMO. Thus, in considering this overall pattern of results PM seems to be a largely independent function with ecologically valid PM tasks showing clear boundary conditions for the applicability of other higher cognitive processes.

In sum, the parallel mediation analyses for indirect age-effects on PM via facets of executive functioning ability, did not reveal shifting, updating, or inhibition, as reliable mediators of low or high demand PM tasks (i.e., generally event and time-based tasks, respectively) on the CAMPROMT, MIST, Virtual Week, and MEMO measures. There was also no significant serial mediated model of age-effects on any PM measure, occurring via the impact of perceptual processing speed on facets of executive functioning. Thus the findings from the empirical investigation in Chapter 5 suggest that individual differences in older adults' PM performance, on high (or low) demand PM tasks, is not invariably associated with individual differences in facets of executive functioning. In particular, when ecologically valid PM measures are used (i.e., the PM and ongoing tasks are not too abstract or unfamiliar) then older adults PM performance in both laboratory and in naturalistic-settings appears to be relatively unaffected by individual differences in facets of executive functioning, processing speed, and often retrospective memory (learning and retention). This is good news for healthy older adults, and resonates with the early, optimistic speculations of Einstein and McDaniel (1990, p. 724) that "prospective memory seems to be an exciting exception to typically found age-related decrements in memory", or executive functioning for that matter.

6.4 Contextualising the Current Findings

6.4.1 Age-PM paradox

The first study to endeavor to generate parallel tasks across laboratory and naturalistic-settings was Rendell and Craik (2000). In their study, Virtual Week was developed to try to resolve the age-PM paradox by having a laboratory paradigm that in many ways resembled daily life. Contrary to their predictions, younger adults continued to outperform older adults in most Virtual Week PM task types. The current research findings using similar parallel PM tasks to

Rendell and Craik (2000) found a similar general pattern in the laboratory, with young outperforming older adults on Virtual Week PM tasks in Experiment 1 and young-old outperforming old-old on Virtual Week PM tasks in Experiment 2. However, as with the finding of Rendell and Craik (2000), there were some nuances in the current findings with performance for both age groups varying across the different PM task types, presumably because of the differing levels of environmental support associated with each task type. In Experiments 1 and 2 (Chapter 4) all age groups performed better on event-based tasks than on time-based task (across settings), suggesting relatively spontaneous processes of retrieval, unaffected by age, were in operation to perform these tasks.

Chapter 4 did not set out to examine focal versus non-focal event-based tasks, or regular versus irregular event and time-based tasks (which correspond with low versus high demand PM tasks in each case respectively). However, an analogous time-based task distinction (time-of-day versus time-interval), assumed to make similar low and high demands on cognitive processes, was explicitly made and tested. Virtual Week *appears* well suited to capture the distinction between time-of-day and time-interval, given the simulated daily narrative that forms the background task of this paradigm. However, while this distinction appears theoretically fruitful, particularly in naturalistic-settings where more experimenter controlled paradigms are being developed and used (Bailey et al., 2010; Niedźwieńska & Barzykowski, 2012), in the present research older adults performed at a similar level on the two time-based task types in Virtual Week. One possible reason for this similar level of performance between the time-of-day and time-interval tasks is that the former type of task (i.e., time-of-day) usually occurs over longer (real) time periods with more opportunity for planning and execution. In naturalistic-settings, time-of-day tasks are also less likely to compete with multiple other PM intentions that need to be executed within a relatively short, real time window. Furthermore, although Virtual Week simulates many of the time-of-day cues that occur in a naturalistic-setting (e.g., by the narrative formed by event cards), it is a formidable methodological challenge to capture the same rich array or constellation of cues that occur in naturalistic-settings that cue individuals to the time of day and concurrent or imminent appointment-like (i.e., time-of-day) PM tasks.

It may also be that the nature of time-interval tasks needs to be further delineated. For example, Niedźwieńska and Barzykowski (2012) used a repeated time-interval task in the laboratory (i.e., writing down the time every 10 minutes) and found no difference between young, middle aged and older adults. However, when the same participants were given a time-of-day task in a naturalistic-setting (i.e., phone the experimenter at a specific pre-arranged time each morning for the next four days) the typical pattern of older adults performing better than young adults emerged. (Interestingly, middle aged adults' performance was intermediate between the two extreme age groups, providing potential clues as to the environmental support that may increase with age in naturalistic-settings; e.g., more routine and structure to use as a framework within which to embed time-of-day tasks that arise irregularly.) The finding by Niedźwieńska and Barzykowski (2012) of no differences between young, middle aged, and older adult age groups on a time-interval task (writing down the time every 10 minutes) in the laboratory suggests that other factors, such as how absorbing the ongoing task is, and/or how much attention is being allocated to each task in a dual (or multi) task situation may also be relevant to PM task performance. It may also be that whereas in Virtual Week the time-interval is relatively short (2 minutes and 4 minutes on the stop clock) a longer time-interval (e.g., 10 minutes) may give older adults more chance to organize themselves in terms of allocating attention strategically. These possibilities of directly comparing different durations of delay in time-interval tasks in Virtual Week with young and older adults await empirical investigation.

When considering the disconnect between older adults' PM performance on different PM task types across settings, there is a temptation to assume that the results from one setting (either laboratory or naturalistic) represent the "real" measure of age-effects on PM. However, this either-or reasoning, or even reification of one PM paradigm (e.g., the classic paradigm), is fallacious and unproductive for developing a richer theoretical understanding of cognitive age-effects on PM. A better approach is to critically consider the conceptually divergent or parallel characteristics of PM tasks across settings. This is what has been attempted in Chapter 4 of the current research project. The results presented in Chapter 4, and reviewed above, suggest that an enhanced understanding of age-effects on PM can be obtained by carrying out empirical investigations that conceptually match PM tasks across settings, *and* focus on subtypes of time-based PM tasks (i.e., time-of-day and time-interval). The arguments and empirical research presented in the current research project do not aim or wish to imply that previous studies in laboratory-settings give an inaccurate picture of age-effects on PM, nor that previous, naturalistic-setting studies provide a distorted or flawed picture due to poor validity arising from lack of experimenter control over key variables. Rather it should be recognized that previous studies have differed in the types of event (e.g., Rendell & Craik, 2000) and time-based PM tasks used in each setting (e.g., Schnitzspahn et al., 2011). In the case of time-based tasks, it can be clearly seen that both types of time-based tasks (i.e., time-of-day and time-interval) have merit as representing *some* PM task types that occur in real life.

In sum, laboratory and naturalistic-setting studies of age-effects on PM are complementary. The age-PM paradox, that the present research project has taken substantial steps towards resolving, should not be viewed as an indication that tasks administered in laboratory-settings lack validity because they are artificial compared to "real" tasks in naturalistic-settings (Ihle et al., 2012; Kliegel et al., 2007; Lee et al., 2018); nor that tasks administered in naturalistic-settings lack validity because they are contaminated and uncontrolled in comparison to well controlled tasks in the laboratory (Phillips et al., 2008). This fallacy can be illustrated. For example, the experiment that led to the coining of the term age-PM paradox (Rendell & Thomson, 1999) included a task conducted over several days in the naturalistic-setting but used an artificial task (pressing a button on a time logging device at set times). In contrast, the Rendell and Craik (2000) study included an hour or so in a laboratory-setting using Virtual Week that involved a representative simulation of PM in daily life in an engaging board game format. The question may then arise as to whether individual differences in executive functioning are relevant for explaining age-effects in all or only some of these different PM paradigms and task types.

6.4.2 Age, executive functioning, and PM

The investigation of how executive functioning relates to age-effects on PM has primarily been the preserve of abstract laboratory paradigms (McDaniel & Einstein, 2007). Two recent studies on individual differences and facets of executive functioning mediating PM, which the current research project findings speak to, are Schnitzspahn et al. (2013) and Zuber et al. (2016). Most similar in aims to those presented in the current research project in Chapter 5, was the investigation of Schnitzspahn et al. (2013). Schnitzspahn et al. (2013) investigated the relative importance of processing speed, shifting, updating, and inhibition, as mediators of PM performance in young and older adults using a version of the classic Einstein and McDaniel (1990) laboratory paradigm. However, in contrast to the current research projects' findings, Schnitzspahn et al. (2013) found that shifting and inhibition (though not updating) were modest but significant predictors, and explained some variability associated with age, on a non-focal event-based PM task.

It is worth considering some potentially important differences in design between the current research project (see Chapter 5) and that of Schnitzspahn et al. (2013). First, unlike the present research project, Schnitzspahn et al. (2013) used two tasks to measure each construct (shifting, updating, and inhibition), and these tasks closely resembled those used by Miyake et al. (2000). The task switch task used in the present research project closely resembles one of the shifting tasks (the semantic category task) used by Schnitzspahn et al. (2013). However, there were some important scoring differences. Schnitzspahn et al. (2013) used only correct responses to calculate a global switch cost between pure and mixed block trials. The present study utilised all data (correct and incorrect latencies) to calculate a rate residual score (Hughes et al., 2014).[20] This score is based on the higher (better) or lower (poorer) than expected rate of responding accurately on switch and non-switch trials, using a linear regression model predicted from overall group performance. This is a more conservative scoring method (*qua* reducing rate of Type 1 errors) and suggests that, if mental set shifting mediates age-effects on high demand PM task types, then the difference in findings regarding shifting between the present study and Schnitzspahn et al. (2013) is a function of the different characteristics of the PM measures used (e.g., the former using more familiar ongoing tasks and cue-action pairs).

In contrast to the shifting task, the updating tasks used by Schnitzspahn et al. (2013) is less similar to the *n*-back updating task used in the present research. The letter-memory task has some similarity to the *n*-back but perhaps not with same sustained amount of attention required,

[20] In the present study the correlation between the rate residual score and the global-switch cost score was $r = .93$, $p < .01$. The correlation between global switch cost score and local switch cost score was $r = .35$, $p < .01$; the rate residual correlation with the local switch cost score was $r = .36$, $p < .01$. Both the global and local switch cost scores did not significantly correlate with any of the PM measures used.

given that trials were separated by short breaks, while the *n*-back presents stimuli in a more inexorable fashion. The inhibition measures also diverge somewhat in that those used by Schnitzspahn and colleagues exploit more naturally habitual responses (e.g., in the anti-saccade task looking in the direction of briefly presented stimuli instead of at the opposite side of a screen), whereas the Go-no go appears to generate a habitual form of responding over time, with many trials. It is possible that these slight differences are inconsequential. However, this is an empirical question, with construct validity favouring Schnitzpahn and colleagues who more closely followed the executive functioning measures used in the factor analysis of Miyake et al. (2000).

However, it is here proposed that the main difference in measures used between the present research project and Schnitzspahn et al. (2013) is the PM paradigm itself. In comparison to the non-focal event-based tasks embedded within lexical decision making ongoing tasks used by Schnitzspahn and colleagues, the present research project used a diverse array of PM measures in which, arguably, the ongoing and PM tasks had high ecological validity. The ongoing tasks in Schnitzspahn et al. (2013) in many ways resembled a shifting paradigm, i.e., lexical decisions with arrow key button presses to indicate whether a word on the left or right contained more syllables, or was in the same or a different semantic category. Similarly the PM task in the first ongoing task was to press the space bar if a verb appeared, while in the latter it was to press the space bar if the word appeared in a particular colour (reminiscent of the Stroop test of inhibition). Given these family resemblances between the ongoing tasks and traditional measures of executive functioning, it is perhaps not surprising that Schnitzpahn and colleagues found shifting and inhibition were significant predictors of PM performance on these tasks (cf. Marsh & Hicks, 1998). What is surprising, and converges with the present research findings of executive functioning across diverse PM outcome measures consistently accounting for only 0–3% of variance in PM performance, is that in the hierarchical model presented by Schnitzspahn et al. (2013, p. 1550) the inclusion of shifting, updating and inhibition only accounted for an additional 4% of variability on a putatively high demand PM task.

In considering the putative low and high demands of different types of PM tasks, and whether facets of executive functioning are primarily related to the latter type and not the former (as the multiprocess framework proposes), it is useful to consider the measures and findings of Zuber et al. (2016). Many of the measures used by Zuber et al. (2016) exactly corresponded to those used in the present research project: the same version of the 2-back (*n*-back) task (used as both a measure of updating, and as the ongoing task); Go-no go with letters as stimuli (inhibition); and task-switch task, judgments of number magnitude and parity (shifting). However, the designs diverged in the PM measure used. In the case of Zuber and colleagues focal and non-focal event-based tasks were used, while in the present research project both event (though not focal versus non-focal) and time-based tasks (varying in putative demands) within multiple paradigms were incorporated. There was also divergence in the age range of the sample used. Zuber and colleagues used a young to young-old sample (20–68 years), while the present research project used a young-old to old-old sample (60–87 years). Thus, it is possible that both studies were affected by a restricted age range. However, in either case this does not seem to be a compelling argument for a caveat or limitation, given it is widely recognized that most adult cognitive change and decline occurs from 60 years of age (Kliegel, MacKinlay, et al., 2008). Furthermore, in the present research project it could be argued that young-old should be particularly contrasted in terms of performance and individual differences in facets of executive functioning ability relative to old-old adults if cognitive decline accelerates in the last phases of the lifespan.

A curious convergence in findings between Zuber et al. (2016) and the present research project was found in relation to inhibition being a significant predictor of putatively low (focal event-based, and naturalistic-setting event-based) PM tasks. However, at least in relation to the present research project, this finding should be interpreted cautiously given the number of inferential analyses run (and thus the risk of Type 1 errors) and the isolated nature of the finding.

As noted when comparing the results of the present research project with those of Schnitzspahn et al. (2013), so to with Zuber et al. (2016) the difference in the ongoing and PM tasks seems particularly relevant in explaining the differences in findings. There seems to be a trend when investigating the role of facets of executive functioning to use a version of the classic PM paradigm (Einstein & McDaniel, 1990) with a traditional measure of executive functioning as the ongoing task (Joly-Burra et al., 2018; Marsh, Hicks, & Hancock, 2000), which may bias estimates of the relative importance of facets of executive functioning in daily life PM tasks for older adults. In comparison, the present research project is one of the few to use relatively more ecologically valid ongoing and PM tasks when considering the role of executive functioning mediating age-effects on PM (Azzopardi et al., 2015; Rose et al., 2010). The reason for qualifying the ecological validity of the present research project's ongoing and PM tasks as *relatively* better, in this respect, than the classic paradigm, is to acknowledge the potential objection that these types of task are still at a considerable remove from actual ongoing and PM tasks older adults perform in daily life. For example, in the two clinical measures, the MIST and the CAMPROMT, the word puzzles that form the ongoing task are not particularly likely to be representative of the ongoing tasks older adults are involved in. From Chapter 4, it was possible to get a small glimpse into the actual ongoing tasks of older adults (who remembered to answer the scheduled and pop up quizzes). It can be speculated that those older adults who happened to be engaged in cross word or Sudoku puzzles prior to completing the PM task on the MEMO would likely classify it as a leisure activity, while those who were engaged in chores were probably more physically active than when doing the ongoing tasks of the MIST and CAMPROMT. There is also the related issue of differences in self-pacing (e.g., no participants worked furiously to find all the words in the MIST), enabling the freeing up of attentional resources to complete the required PM tasks. Thus, only when executive functioning is particularly taxed by an ongoing task, that demands similar resources to those required on the PM task (Marsh & Hicks, 1998) may the right conditions obtain for individual differences in executive functioning to be relevant.

A more direct comparison, in terms of ongoing and PM tasks, can be made between the present research project and the study by Kamat et al. (2014), who used the MIST with healthy older adults. In Kamat et al. (2014) older adults' performance on the putatively high demand time-based PM tasks was diminished in comparison with the putatively low demand event-based tasks. Furthermore, Kamat and colleagues found that a composite measure of executive functioning was correlated with total and time-based PM scores (though not event-based scores). In contrast, the MIST results presented from the present research project showed that processing speed was correlated with total and time-based scores, and that updating was associated with total score. However, when entered in a two stage regression model, age was the only significant predictor at both stages and only for total MIST score. The difference between the present findings and those of Kamat may reflect the different measures of executive functioning used (i.e., a global measure versus facets of executive functioning). Partially consistent with Kamat and the differential putative demands of the two types of PM tasks, neither age or any facet of executive functioning was a significant predictor of performance in the regression model used in the present study (though age was negatively correlated with event-based scores, $r = .21$, $p < .05$).

6.4.3 Theoretical Implications

> *If people are deprived of what Gibson might have called the rich and convenient "temporal affordances" of their normal routines, they can only schedule their time-keeping with respect to precise external or fallible internal clocks, or by monitoring some regular metabolic process such as their own heart rate or breathing. They have no other options because, in this universe, the passage of time is only manifest as successive, perceptible changes in physical processes.*
>
> \- Rabbitt (1996, pp. 240-241)

The theoretical implications that can be drawn from the empirical findings in this thesis vary depending on the theory under consideration. In terms of the multiprocess framework, the current findings are largely consistent, with the main implication being an extension needed to the consideration of time-of-day and time-interval tasks being analogous to focal and non-focal event-based tasks in terms of relatively spontaneous retrieval processes versus more effortful, strategic retrieval processes. It should be noted that time-based tasks were not the primary focus of McDaniel and Einstein (2000) when developing the multiprocess framework. However, they do note that: "Aspects of the theoretical work presented herein may also be relevant to time-based tasks" (McDaniel & Einstein, 2000, p. 128). The hypothesis that there is an important distinction between the cognitive processes involved in different time-based tasks that may potentially account for aspects of the age-PM paradox was supported when considering time-of-day and time-interval tasks in a naturalistic-setting (at least in Experiment 1, see Chapter 4). The statistically equivalent performance of time-of-day and time-interval tasks in Virtual Week (Experiment 1 and 2 of Chapter 4) might suggest that there are similar cognitive demands on these tasks per se (consistent with the tendency to see time-based tasks as homogenous). However, drawing the conclusion that time-of-day and time-interval tasks are equivalent in terms of cognitive resources in the laboratory, is not entailed by the similar levels of performance on Virtual Week. Furthermore, it can be seen from the current research project's results that there was a trend towards a substantive distinction between these two types of task when comparing young-old and old-old, which may be partially obscured by potential floor effects. Thus, there is an important indication here of a distinction within time-based tasks that may correspond to the relatively spontaneous and controlled processes postulated by the multiprocess framework to account for apparent discrepancies within empirical findings on event-based laboratory tasks (McDaniel & Einstein, 2000).

In regards to a wider range of theories of PM—i.e., the process model (Kliegel et al., 2002), the self-initiated operations model (Craik, 1986), and the pre-attentional processes model (Smith, 2003)—the current findings in relation to the role of executive functions for age-effects on diverse PM measures has potentially more serious implications. In particular, the scope of these theories is put in question by the consistent finding of the present research project (see Chapter 5) that facets of executive functioning were rarely associated with, did not predict, or mediate age-effects on PM with a wide range of measures with high ecological validity. In some cases, such as on the CAMPROMT and the MEMO in Experiment 2, this may reflect the use of note taking or external aids. It is noteworthy that it was on the laboratory paradigm, Virtual Week, and the MIST clinical measure, which both disallow note taking or external aids that age-effects manifested, though again this was not mediated by the measures of facets of executive functioning used in the present research project.

6.4.4 Methodological Considerations

6.4.4.1 Measuring prospective memory. A signature strength of the current research project was to deploy multiple paradigms in testing age-effects on PM with older adults. However, one paradigm that was not used was the classic Einstein and McDaniel (1990) paradigm. This omission does not reduce the value of the present research or indicate that this paradigm does not shed light on important aspects of ageing and PM. Instead, while acknowledging the importance of continuing research using the classic paradigm, it is here claimed that there is also an equally important need to investigate age-effects on PM using a diverse range of PM tasks that are high in ecological validity. Furthermore, the low number of PM trials in the classic paradigm, which translates into lower reliability (particularly in individual differences research), argues for the need to continue using measures like the ones used in this research project, in which there are multiple PM trials (up to 20 PM trials in the case of the two day version of Virtual Week).

6.4.4.2 Conceptually parallel tasks across settings. Continuing work on how the age-PM paradox functions will require further investigations using conceptually parallel PM tasks across settings. In the PM literature, there is much work on how characteristics of event-based tasks interact with age in the laboratory (Kliegel, Jäger, et al., 2008), and even some in naturalistic-settings (Bailey et al., 2010; Niedźwieńska & Barzykowski, 2012). However, to date there has been relative neglect of characteristics of time-based tasks that might show similar interactions with age. The distinction within time-based tasks proposed in this thesis was between time-of-day and time-interval tasks, the former marked by an array of time-based (e.g., position of the sun) and social cues (e.g., regular meal times), the latter by relatively arbitrary intervals of time, involving PM tasks that unexpectedly occur during the day (e.g., a request to call someone back in 10 minutes; Rose et al., 2015). Clearly this distinction has more relevance in naturalistic-settings where time-of-day cues are abundant. In contrast, developing this time-based task distinction in the laboratory is more challenging due to the relative lack of time-of-day cues. It was argued in this thesis that the best laboratory PM paradigm to capture both types of time-based tasks in the laboratory, is Virtual Week, which simulates at least some of these time-of-day cues. The challenge to develop distinct time-based tasks in the laboratory is in some ways a challenge analogous to measuring event-based task performance using focal and non-focal cues in naturalistic-settings. In regards to event-based tasks, the relatively new MEMO paradigm provides the possibility for meeting the challenge of having reasonable control in verifying event-based task performance in naturalistic-settings, and capturing aspects of event-based cues (e.g., regular versus irregular; short events versus extended events).

6.4.4.3 Real time and virtual time. The construct of time is partially constitutive of the nature of PM tasks. Time must elapse between the forming of an intention and the opportunity to execute it (McDaniel & Einstein, 2007), or even the cue to execute an intention and the opportunity to do so (Einstein et al., 2000). The time delays in naturalistic-settings for PM tasks are frequently in the order of hours or even days. In contrast the time delays in the laboratory or clinic for tests of PM are frequently in the order of minutes. This disconnect between types of time-based tasks measured in laboratory and naturalistic-settings is understandable from a practical, methodological point of view. There is a formidable (perhaps impossible) challenge in developing time-of-day time-based tasks in the laboratory. However, it is important to attempt this task, and one methodology for doing this is using simulated or virtual times of day, as in

Virtual Week. The processes of time-perception and time-monitoring, and the nature of the age-related decline in these processes for older adults, is not well established (Mioni, Cardullo, Ciavarelli, & Stablum, 2019). The internal "calibrating" of virtual time (which changes as a function of the participant moving the token around the board in Virtual Week) will inevitably involve different processes to that used for the internal calibration of actual time (as when a stop clock is used; Ceci & Bronfenbrenner, 1985), and again for time-of-day, in which broader windows and more rhythmic, biological processes are likely to be involved (Rabbitt, 1996). These are important questions that require the development of existing methodologies such as Virtual Week and MEMO, and perhaps the development of other, new paradigms.

6.4.4.4 The variety of executive functioning measures. The Miyake et al. (2000) model of separable executive functions (shifting, updating, and inhibition) has been important in recent developments within the cognitive ageing and PM field. One motivation for the factor analysis undertaken by Miyake et al. (2000) was the great diversity of putative measures of executive functioning. The measures of shifting, updating, and inhibition used in the present study were the same as many of the measures used by Zuber et al. (2016) in their analysis of the factor structure of focal and non-focal event-based tasks using the classic PM paradigm. However, these measures of facets of executive functioning were not the same as those used by Schnitzspahn et al. (2013), who used the same measures as Miyake et al. (2000). Although, similar findings for shifting predicting performance on high demand (i.e., non-focal) event based tasks was found by both Schnitzspahn et al. (2013) and Zuber et al. (2016), these two studies diverged in relation to the role of inhibition. Schnitzspahn et al. (2013) found inhibition was related to high demand (non-focal) event-based tasks, while Zuber et al. (2016) found that inhibition was related to a low demand (i.e., focal) event-based tasks. These discrepant results, compounded by the general lack of association between facets of executive functioning with diverse measures of ecologically valid measures of putatively low and high demand PM tasks in the current research project, suggest that careful consideration of the executive functioning measures used, and how they might relate to diverse PM tasks, is warranted for future studies in the cognitive ageing and PM field.

6.5 Strengths and Limitations

There are a number of key strengths of the current research project investigation into cognitive ageing and PM. First, there was the use of conceptually parallel PM task types across settings to shed light on the potential methodological mechanisms driving the age-PM paradox. This clarified the critical importance of levels of environmental support for different PM task types in understanding age-effects in both laboratory and naturalistic-settings. Chapter 4 also brought out a critical distinction between time-based tasks that may prove key to the solution of the age-PM paradox. The ability of healthy older adults to compensate for cognitive decline in meeting high demand PM tasks (i.e., time-interval tasks) was also demonstrated. The other major strength of the current research project was the use of multiple, diverse measures of PM (laboratory, clinical, and naturalistic) in investigating the potential mediating role of executive functions, processing speed, and facets of retrospective memory (learning and retention). This led to a healthy critical reappraisal of the extent to which executive processes mediate age-effects on a wide range of PM tasks. Thus, the current research plays a key role in helping to delineate possible boundary conditions for the role of executive functions, and further the current literature by carefully considering specific facets of executive functions and the role they may play on different PM tasks.

There were some limitations of the current research project. These included a potential age range restriction in the sample for the investigation of the role of facets of executive functions mediating age-effects. It was argued that this was not a major limitation as there is still expected to be significant cognitive decline within an older adult cohort. Nevertheless, increasing the age range to include middle aged or young adults may have led to some of the findings that trended towards significance (e.g., the role of inhibition), showing a greater effect. The highly educated (the mean estimated IQ was well above average for the population) and healthy older sample used in the current research project may also constitute a similar range restriction.

6.6 The Unique Contributions of the Current Research

A key unique contributions of the current research has been to show that *there is no real age-PM paradox* once conceptually parallel PM tasks are used in investigating young, young-old, and old-old adults PM performance across settings. Related to this accounting for the age-PM paradox was the development of the distinction between time-of-day and time-interval tasks. In contributing to the literature on the age-PM paradox, the current research project has drawn attention to the robust second aspect of the paradox, i.e., young-old outperforming old-old in the laboratory but showing no age differences in naturalistic-settings. This is an important empirical contribution in itself, as many previous studies have only used extreme age group contrasts, and neglected potential nuances within older adult cohorts. The finding that older adults could also compensate for deficits in PM shown in the laboratory on high demand PM tasks (e.g., time-interval) when permitted to use any strategy they wished in daily life shows that old-old are not as vulnerable to forgetting as is often assumed.

The other major unique contribution to the literature was the finding that PM functioning is a process in its own right, in some cases affected by ageing but also often largely independent of age-effects and individual differences in facets of executive functioning. PM is not therefore a proxy for other cognitive processes, such as a combination of executive and retrospective memory processes. While it was acknowledged that the continuing research on the boundary conditions for executive functioning mediating age-effects on PM is important, the current research project shows that PM is a richer and more complex construct than is often assumed when using the classic laboratory paradigm for studying age-effects on PM. It might be thought that a caveat ought to be included for the current research project not finding age-effects on high demand PM tasks being mediated by facets of executive functioning. This caveat might point to the generally low reliability found for some of the PM measures used. However, in defense of the applicability of the current research project's findings it can be noted that the classic PM paradigm, which typically includes only 1–3 PM tasks, also tends to have low reliability (Kelemen et al., 2006).

6.7 Future Directions

There are some clear directions for future research which have been identified as a result of the findings of the current research project. In regards to the age-PM paradox, there is the need to continue to develop paradigms, such as the MEMO, to learn more about contextual aspects of successful PM performance in naturalistic-settings (i.e., what locations and activities participants are embedded in while maintaining and executing PM tasks). There are also many dimensions along which both Virtual Week and MEMO can be manipulated in order to test various hypotheses about age-effects on PM. For example, in the present research project both Virtual Week and MEMO only occurred over a short number of days (i.e., two test days in Virtual Week, and three test days in MEMO). To explore the possible modulating effect of regular and irregular PM tasks it may be necessary to extend the time frame, e.g., up to a week, to explore these other dimensions of PM tasks. The MEMO may also be developed to test focal and non-focal event-based tasks in a naturalistic-setting, perhaps by using the morning notifications to embed some survey questions and an event-based cue (see Bailey et al., 2010). Further research and data gathering on potential strategies used by older adults to remember PM tasks could also be captured using further developed versions of the MEMO paradigm.

In terms of the role of facets of executive functioning in mediating age-effects on high (and low) demand PM tasks, there are multiple future directions which present themselves from the research findings that these did not appear to play a role on the PM measures used in this study. One obvious comparison that needs to be directly made within the one study is that between the classic laboratory paradigm and other PM measures with more naturalistic ongoing tasks (though, as noted above, the classic paradigm is limited in applicability to correlational and individual differences research).

6.8 Conclusions

The importance of continuing to empirically investigate age related changes in PM is obvious given the ageing demographics of societies around the world. However, how this investigation is carried out largely depends on the theories that inform this research. The present research project has made a contribution to the theoretical development of age-effects on PM by highlighting a frequently overlooked distinction within time-based PM tasks (i.e., time-of-day versus time-interval). This distinction sheds light on mechanisms driving both aspects of the age-PM paradox (i.e., young adults outperforming older adults in laboratory-settings but vice versa in naturalistic-settings; and young-old outperforming old-old in the laboratory but showing no difference in naturalistic-settings). The need for careful methodological considerations when comparing PM tasks across settings, and being circumspect in generalising findings from one paradigm to others, has also been demonstrated. In particular, the role of executive functioning in accounting for age related declines in PM needs to be carefully considered in terms of boundary conditions for settings and paradigms used, and other influences on PM for older adults need to be considered and explored. It appears that as paradigms include more naturalistic components (e.g., ongoing task or PM tasks) the relevance of executive functioning diminishes.

Appendices

Appendix A–1: Original Parent Study Ethics Approval

From:	Pratigya Pozniak on behalf of Res Ethics
To:	Peter Rendell; Colleen Doyle; Nathan Rose
Cc:	Res Ethics; Skye McLennan; Denise Rose; Biljana Vukovic; erin.sinclair@myacu.edu.au; sarah.gatt@myacu.edu.au
Subject:	2015-259H Ethics application approved!
Date:	Friday, 29 January 2016 1:52:11 PM

Dear Applicant,

Principal Investigator: Prof Peter Rendell Prof Colleen Doyle, A/Prof Julie Henry, Dr Nathan Rose,
Co-Investigator: Prof Matthias Kliegel, Dr Skye McLennan,
Student Researcher: Erin Rose Sinclair (HDR Student), Sarah Lorraine Gatt (HDR Student)
Research Assistant: Mrs Denise Rose, Biljana Vukovic
Ethics Register Number: 2015-259H
Project Title: Acting with the future in mind
Risk Level: More than Low Risk
Date Approved: 29/01/2016
Ethics Clearance End Date: 31/12/2018

This email is to advise that your application has been reviewed by the Australian Catholic University's Human Research Ethics Committee and confirmed as meeting the requirements of the National Statement on Ethical Conduct in Human Research.

Appendix A–2: Modification of Parent Study To Include Thesis Author

HREC Register No.	2015-259H
Approval End Date	29/01/2016
Principal Investigator/Supervisor	Prof Peter Rendell
Student Researcher (if applicable)	Erin Rose Sinclair, Sarah Lorraine Gatt, Joshua Ranger, Bianca Stephenson-Gromer, and Simon Haines
Protocol Title	Acting with the future in mind

Proposed modifications to the Protocol:

NOTE: Such modifications may include changes in the aim, procedures or direction of the protocol, sources or manner of recruitment of subjects, number or age of participants, changes to the questionnaire, survey instruments, Letter(s) to the Participants, Consent Forms or changes to personnel.

State what the proposed modification is:

We request four minor changes (1) additional investigators be included in the project (2) the addition of a short questionnaire (State-Trait Anxiety Inventory), (3) a modification to the recruitment flyer and (4) remove upper age restriction of participants. All other aspects of the study remain unchanged .

REQUESTED MODIFICATIONS

1. Additional Investigators.

We request approval for two research assistants, one PhD student, and two honours students to be added to the project. The research assistants are Rachel Pelly and Nicholas Koleits. Rachel has completed her honours degree with ACU at the end of 2015 and has been working with the university as a research assistant for eight months. Nicholas completed his undergraduate and post graduate diploma in psychology at ACU in 2014 and has been working on a variety of projects as a research assistant with the university for the past year. Both research assistants have worked with Cognition and Emotion Research Centre at Australian Catholic University, Melbourne and bring valuable skills and knowledge to the project. They will be involved in developing the research protocol, participant recruitment, and data collection and storage. The PhD student is Simon Haines. Simon completed his honours degree with ACU in 2015 and will be involved in the study as part of his Ph.D. project. Simon is being supervised by Professor Peter Rendell and Dr Skye McLennan (who are listed as PI and CI on this project).

Appendix B: Information Letter

PARTICIPANT INFORMATION LETTER

PROJECT TITLE: MemoryTrain
PRINCIPAL INVESTIGATOR: Prof Peter Rendell (ACU)
Chief-investigators: Prof Colleen Doyle (Villa Maria Catholic Homes), Assoc Prof Julie Henry,(University of Queensland)
Prof Matthias Kliegel (University of Geneva & ACU), Dr Nathan Rose (ACU), Dr Skye McLennan (ACU)
STUDENT RESEARCHERS: Erin Sinclair, Sarah Gatt, Simon Haines and Joshua Ranger
STUDENT'S DEGREE: Bachelor of Psychological Science (Hons), Masters of Psychology (ACU) and Ph.D.
Research Assistants: Ms Biljana Vukovic, Mrs Denise Rose, Ms. Rachel Pelly and Mr. Nicholas Koleits

Dear Participant,

You are invited to participate in the research project described below.

What is the project about?
This study investigates prospective memory - the act of remembering future intentions. Prospective memory is involved in many daily activities that are critical for the maintenance of independence in late adulthood, such as remembering to take medication and turn off appliances.

This project will aim to establish whether it is possible to enhance the prospective memory function of healthy older adults. All participants are aged between 60 and 89 years old, you and will be assigned to one of four different memory training programs and take part in an introductory session.

There will be six weeks of computer based training, which has been automated so you can complete it at home. At the beginning of the study, after the six week training program and three months after the study, you will be asked to complete some memory exercises, puzzles and questionnaires to determine which type of training is most effective.

The results will have direct and immediate implications for clarifying how prospective memory function can be optimised in late adulthood.

Who is undertaking the project?
The project will be undertaken by Professor Peter Rendell (School of Psychology, ACU), Director of the Cognition and Emotion Research Centre and a registered psychologist. This project is part of an ongoing research investigating prospective memory and aging with co-investigators: Professor Colleen Doyle (Villa Maria Catholic Homes & ACU), Professor Matthias Kliegel (University of Geneva & ACU), Associate Professor Julie Henry (University of Queensland), Dr Nathan Rose, and Dr Skye McLennan (School of Psychology, ACU). This major project is funded by an Australian Research Council (ARC) Linkage grant with partner Villa Maria Catholic Homes that is titled: Acting with the Future in Mind. In addition, students - Erin Sinclair and Sarah Gatt – will be participating in this research as part of their Masters of Psychology degree at Australian Catholic University.

Are there any risks associated with participating in this project?
During the course of the research, some personal information will be collected, for example your name and phone number. Only researchers will have access to this information. To ensure your personal information is secure, it will be stored in a locked cabinet in the primary researcher's office at the

Australian Catholic University. There are no expected psychological risks associated with the program, however if at any time you feel uncomfortable you can choose to leave the study.

What will I be asked to do?

You will firstly complete a phone screening session, whereby, you will be asked to complete a series of questions over the telephone to ensure you are eligible to participate in the study. If you are eligible to participate in the study, you can arrange a time to undertake the six week training session and the three testing session, one pre training and two post training.

The first session is the pre training test session where you will be asked some background questions about age, gender and general health. You will also be asked to complete several short pen and paper puzzle activities, as well as, computer-based activities. One task requires that you read words aloud. We will need to record your responses for this task via audiotape in order to ensure accurate scoring.

Within a few weeks of the testing session, you will then be asked complete the six week training program. Prior to commencing the training program, a research assistant will meet with you to fully explain and set up the computer based training. The training program occurs at your home, in your own time, on a computer. If you do not own a computer, we can lend you one for the duration of the study. The training involves about two hours of activities per week, either 4 weekly sessions (of 25 to 30 minutes) of the computerised training, or a weekly one hour session of computerised training with up to a one hour of homework activities. The training days and times are determined by you. You will have telephone support from the researcher to assist you with any issue or questions as well as to check your progress.

Some participants will be asked to complete a mobile phone task. You will be given the option to use your own phone or one of our mobile phones which functions as a multiple choice questionnaire completed in real time. You will be prompted about three times a day for up to three days, to answer a few questions. You will also be asked to take photos with the phone to show tasks you have completed. Only those who are comfortable with the mobile phone will be asked to participate. Participants will be given a briefing during testing and a chance to practice the first questionnaire. You will also be provided with a contact phone number in case assistance is needed during the six day period.

A couple of weeks after the training is completed, a follow-up testing session occurs, which is very similar to the first testing session. Participants completing the phone task will have a chance to transfer data and/or return the study phone. In addition, the final follow up testing session will occur three months after the training is completed.

How much time will the project take?

The telephone screening session will take approximately 20 minutes. The first testing session will take approximately 2.5 – 3 hours. The second session should take approximately 2 hours. As described above, the training will require two hours a week, over six weeks, with either 4 separate sessions of 30 minutes or one hour tutorial session with one hour of homework tasks. Both follow up testing sessions are expected to take 2 hours.

What are the benefits of the research project?

There is the potential for improvements in your memory performance; however, there is also the potential for no improvements to occur. Although you may not benefit from this research personally, we hope that the results from the research study will benefit society by helping the advancement of theoretical understanding of memory training.

Can I withdraw from the study?

Participation in this study is completely voluntary. You are not under any obligation to participate. If you agree to participate, you can withdraw from the study at any time without adverse consequences.

Will anyone else know the results of the project?
Individual results of the project will remain confidential. During the project, all participants will be given a code and names will be stored separately and not retained with the data. The results of the study will be published in the form of a masters thesis and we also plan to report the findings at a conference and/or in a scientific journal. In all reports of the results, only group results will be reported; hence, individual participants will not be able to be identified in any reports.

Will I be able to find out the results of the project?
If selecting the feedback option in the consent form and/or upon written request, results of the completed study can be made available to you twelve months after the study has been completed (2019). Please address correspondence to Professor Peter Rendell, Locked Bag 4115, Fitzroy MDC, VIC 3065.

Who do I contact if I have questions about the project?
Any questions regarding this project can be directed to the Principal Investigator: Professor Peter Rendell in the School of Psychology, St. Patrick's Campus (Australian Catholic University, Level 5, The Daniel Mannix Building, Young Street, Fitzroy 3065, phone 03 9953 3126)

What if I have a complaint or any concerns?
The study has been reviewed by the Human Research Ethics Committee at Australian Catholic University (review number 2014 xxxx). If you have any complaints or concerns about the conduct of the project, you may write to the Manager of the Human Research Ethics Committee care of the Office of the Deputy Vice Chancellor (Research).

Manager, Ethics
c/o Office of the Deputy Vice Chancellor (Research)
Australian Catholic University
North Sydney Campus
PO Box 968
NORTH SYDNEY, NSW 2059
Ph.: 02 9739 2519
Fax: 02 9739 2870
Email: resethics.manager@acu.edu.au

Any complaint or concern will be treated in confidence and fully investigated. You will be informed of the outcome.

I want to participate! How do I sign up?

If you are willing to participate please sign the attached informed consent forms. You should sign both copies of the consent form and keep one copy for your records and return the other copy to the staff supervisor. Your support for the research project will be most appreciated.

Yours sincerely,

[redacted]

Professor
Peter Rendell
Principal investigator

Appendix C–1: Evidence of submission of Oxford article

Oxford Research Encyclopedia of Psychology <onbehalfof@manuscriptcentral.com> Fri, Jun 22, 2018, 1:26 PM
to Simon.Haines80

Dear Mr. Haines,

The article titled Prospective Memory and Cognitive Aging has been submitted by Mr. Simon Haines to Oxford Research Encyclopedia of Psychology.

As you are listed as a co-author, you can check the status of this article or other articles you have authored/co-authored. The online peer-review system, ScholarOne, automatically creates a user account for you. Your username and password for your account are:

Site URL: https://mc.manuscriptcentral.com/orepsy
USER NAME: Simon.Haines80@gmail.com
PASSWORD: Clicking the link below will take you directly to the option for setting your permanent password.

https://mc.manuscriptcentral.com/orepsy?URL_MASK=31bd1228c3a0495298043120c7d42d0b

When you log in for the first time, you will be asked to complete your full postal address, telephone, and fax number. You will also be asked to select a number of keywords describing your particular area(s) of expertise.

Thank you for your participation.

Sincerely,
Rosa

Rosa Kneller
Development Editor

Appendix C–2: Evidence of Acceptance and Publication of Oxford Article

From: **ORE Editorial** <oreeditorial@oup.com>
Date: Thu, 28 Feb 2019, 08:56
Subject: ORE Psychology: Congratulations on Your Published Article!
To: Simon.Haines80@gmail.com <Simon.Haines80@gmail.com>

Dear Mr. Haines,

I am delighted to announce that we have just published your article "Prospective Memory and Cognitive Aging" to the *Oxford Research Encyclopedia of Psychology* as part of our February 2019 update. On behalf of the Press, I want to congratulate you and thank you for your contribution; your continued support ensures that the ORE program will remain a dynamic, authoritative resource for years to come.

We invite you to view your article as well as the other articles recently published to the *ORE of Psychology*.

Although you may already have access through your university or institution, we are happy to provide you with temporary gratis access to all subject areas of the Oxford Research Encyclopedias. Please feel free to use the following login details:

OXFORD RESEARCH ENCYCLOPEDIAS

About Contact Us Recently Published Guided Tour Subscriber Services Help

Personal Profile: **Sign in** or **Create**

PSYCHOLOGY

Oxford Research Encyclopedias Psychology

Search

Search within subject: Select ...

Browse by Subfield My Content (1) My Searches (1)

Affective Science
Biological Foundations of Psychology
Clinical Psychology: Disorders and Therapies
Cognitive Psychology/Neuroscience
Developmental Psychology
Educational/School Psychology
Forensic Psychology
Health Psychology
History and Systems of Psychology
Individual Differences
Industrial and Organizational Psychology
Methods and Approaches in Psychology
Neuropsychology
Personality
Psychology and Other Disciplines
Social Psychology
Sport Psychology

Browse All

Back to Results

Subscriber sign in

Username

Password

Sign in Forgot password? Don't have an account?

Sign in via your Institution

Sign in with your library card

Sign in

Prospective Memory and Cognitive Aging

Simon J. Haines, Jill Talley Shelton, Julie D. Henry, Gill Terrett, Thomas Vorwerk, and Peter G. Rendell

Subject: Clinical Psychology: Disorders and Therapies, Cognitive Psychology/Neuroscience

Online Publication Date: Feb 2019 DOI: 10.1093/acrefore/9780190236557.013.381

[-] Summary and Keywords

Tasks that involve remembering to carry out future intentions (such as remembering to attend an appointment), and the cognitive processes that enable the completion of such tasks (such as planning), are referred to as prospective memory (PM). PM is important for promoting quality of life across many domains.

Appendix D: Evidence of Submission of JEP:G Article

Journal of Experimental Psychology: General

Prospective memory: Using parallel tasks to understand contrasting patterns of age differences seen in laboratory versus naturalistic-settings

--Manuscript Draft--

Manuscript Number:	
Full Title:	Prospective memory: Using parallel tasks to understand contrasting patterns of age differences seen in laboratory versus naturalistic-settings
Abstract:	Prospective memory (PM) affords autonomy and quality of life in older age. Prior PM research shows paradoxical findings whereby young adults outperform older adults in laboratory-settings but the reverse is found naturalistic-settings. Moreover, young-old outperform old-old adults in laboratory-settings, but show no age differences in naturalistic-settings. However, task characteristics have differed between laboratory and naturalistic-setting studies. Thus, the paradox of contrasting PM performance

Submission Confirmation for Journal of Experimental Psychology: General - [EMID:4991525bbc0927ea]

Inbox ×

Journal of Experimental Psychology: General <em@editorialmanager.com>

to me

10:11 PM (44 minutes ago)
Mar 15, 2019, 10:11 PM

Dear Mr Haines,

Your submission "Prospective memory: Using parallel tasks to understand contrasting patterns of age differences seen in laboratory versus naturalistic-settings" has been received by Journal of Experimental Psychology: General.

You will be able to check on the progress of your submission by logging on to Editorial Manager as an author. The URL is https://www.editorialmanager.com/xge/.

Your manuscript will be given a reference number once an Editor has been assigned.

Please note that you may also confirm or Authenticate your ORCID iD by clicking here https://www.editorialmanager.com/xge/l.asp?i=233399&l=2BNK82TY.

APA asks that you please take a moment to give us your feedback on the submission process, by completing a short survey, available at http://goo.gl/forms/vKXxocF4Jk.

Best regards,
Editorial Office
Journal of Experimental Psychology: General

References

Aberle, I., Rendell, P. G., Rose, N. S., McDaniel, M. A., & Kliegel, M. (2010). The age prospective memory paradox: Young adults may not give their best outside of the lab. *Developmental Psychology, 46,* 1444–1453. doi: 10.1037/a0020718

Altgassen, M., Kliegel, M., Brandimonte, M., & Filippello, P. (2010). Are older adults more social than younger adults? Social importance increases older adults' prospective memory performance. *Aging, Neuropsychology, and Cognition, 17*, 312-328. doi: 10.1080/13825580903281308

Azzopardi, B., Juhel, J., & Auffray, C. (2015). Aging and performance on laboratory and naturalistic prospective memory tasks: The mediating role of executive flexibility and retrospective memory. *Intelligence*, *52*, 24-35. doi: 10.1016/j.intell.2015.06.0070160-2896

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology: Section A, 49*, 5-28. doi: 10.1080/713755608

Bailey, P. E., Henry, J. D., Rendell, P. G., Phillips, L. H., & Kliegel, M. (2010). Dismantling the “age–prospective memory paradox”: The classic laboratory paradigm simulated in a naturalistic setting. *The Quarterly Journal of Experimental Psychology, 63*, 646-652. doi: 10.1080/17470210903521797

Ballhausen, N., Schnitzspahn, K. M., Horn, S. S., & Kliegel, M. (2017). The interplay of intention maintenance and cue monitoring in younger and older adults’ prospective memory. *Memory and Cognition, 45*, 1113-1125. doi: 10.3758/s13421-017-0720-5

Benedict, R. H. B., Schretlen, D., Groninger, L., & Brandt, J. (1998). The Hopkins Verbal Learning Test-Revised: Normative data and analysis of interform and test-retest reliability. *Clinical Neuropsychologist*, *12*, 43-55. doi: 10.1076/clin.12.1.43.1726

Bisiacchi, P. S., Tarantino, V., & Ciccola, A. (2008). Aging and prospective memory: The role of working memory and monitoring processes. *Aging Clinical and Experimental Research, 20*, 569-577. doi: 10.1007/BF03324886

Brandimonte, M. A., Einstein, G. O., & McDaniel, M. A. (1996). *Prospective memory: Theory and applications*. Mahwah, NJ: Lawrence Erlbaum.

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *Clinical Neuropsychologist, 5*, 125-142. doi: 10.1080/13854049108403297

Brandt, J., & Benedict, R. H. B. (2001). *Hopkins Verbal Learning Test-Revised: Professional manual*. Lutz, FL: PAR.

Braver, T. S., & Barch, D. M. (2002). A theory of cognitive control, aging cognition, and neuromodulation. *Neuroscience & Biobehavioral Reviews, 26*, 809-817.

Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., . . . Mumenthaler, M. S. (2001). Context processing in older adults: evidence for a theory relating cognitive control to neurobiology in healthy aging. *Journal of Experimental Psychology: General, 130*, 746.

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39*, 713-726. doi: 10.1016/S0896-6273(03)00466-5

Braver, T. S., Satpute, A. B., Rush, B. K., Racine, C. A., & Barch, D. M. (2005). Context Processing and Context Maintenance in Healthy Aging and Early Stage Dementia of the Alzheimer's Type. *Psychology and Aging, 20*, 33-46. doi: 10.1037/0882-7974.20.1.33

Breneiser, J. E., & Mcdaniel, M. A. (2006). Discrepancy processes in prospective memory retrieval. *Psychonomic Bulletin and Review, 13*, 837-841. doi: 10.3758/bf03194006

Brewer, G. A., Knight, J. B., Marsh, R. L., & Unsworth, N. (2010). Individual differences in event-based prospective memory: Evidence for multiple processes supporting cue detection. *Memory and Cognition, 38*, 304-311. doi: 0.3758/MC.38.3.304

Brom, S. S., & Kliegel, M. (2014). Improving everyday prospective memory performance in older adults: Comparing cognitive process and strategy training. *Psychology and Aging, 29*, 744. doi: 10.1037/a0037181

Browning, C. A., Harris, C. B., Van Bergen, P., Barnier, A. J., & Rendell, P. G. (2018). Collaboration and prospective memory: Comparing nominal and collaborative group performance in strangers and couples. *Memory, 26*, 1206-1219. doi: 10.1080/09658211.2018.1433215

Burgess, P. W., Quayle, A., & Frith, C. D. (2001). Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia, 39*, 545-555. doi: https://doi.org/10.1016/S0028-3932(00)00149-4

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist, 54*, 165-181. doi: 10.1037/0003-066X.54.3.165

Cavuoto, M. G., Ong, B., Pike, K. E., Nicholas, C. L., & Kinsella, G. J. (2015). Naturalistic prospective memory in older adults: Predictors of performance on a habitual task. *Neuropsychological Rehabilitation, 27*, 1-15. doi: 10.1080/09602011.2015.1074590

Ceci, S. J., & Bronfenbrenner, U. (1985). "Don't forget to take the cupcakes out of the oven": Prospective memory, strategic time-monitoring, and context. *Child Development, 56*, 152-164. doi: 10.2307/1130182

Chen, X.-j., Wang, Y., Liu, L.., Cui, J., Gan, M., Shum, D. H., & Chan, R. C. (2015). The effect of implementation intention on prospective memory: A systematic and meta-analytic review. *Psychiatry Research, 226*, 14-22. doi: 10.1016/j.psychres.2015.01.011

Cherry, K. E., & Lecompte, D. C. (1999). Age and individual differences influence prospective memory. *Psychology and Aging, 14*, 60-76. doi: 10.1037/0882-7974.14.1.60

Cherry, K. E., Martin, R. C., Simmons-D'Gerolamo, S. S., Pinkston, J. B., Griffing, A., & Drew Gouvier, W. (2001). Prospective remembering in younger and older adults: Role of the prospective cue. *Memory, 9*, 177-193. doi: 10.1080/09658210143000092

Cohen, A.-L., Gordon, A., Jaudas, A., Hefer, C., & Dreisbach, G. (2017). Let it go: The flexible engagement and disengagement of monitoring processes in a non-focal prospective memory task. *Psychological Research, 81*, 366-377. doi: 10.1007/s00426-016-0744-7

Costa, A., Caltagirone, C., & Carlesimo, G. A. (2017). Prospective memory functioning in individuals with Parkinson's Disease: A systematic review. *The Clinical Neuropsychologist, 32*, 1-23. doi: 10.1080/13854046.2017.1407457

Craik, F. I. (1986). A functional account of age differences in memory. In F. Klix & H. Heagendorf (Eds.), *Human memory and cognitive capabilities*. Amsterdam, Netherlands: Elsevier.

Craik, F. I., & Bialystok, E. (2006). Planning and task management in older adults: Cooking breakfast. *Memory and Cognition, 34*, 1236-1249. doi: 10.3758/bf03193268

Craik, F. I., & Kerr, S. A. (1996). Prospective memory, aging, and lapses of intention. In M. Brandimonte, G. O. Einstein & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 227-237). Mahwah, NJ: Lawrence Erlbaum.

d'Ydewalle, G., Luwel, K., & Brunfaut, E. (1999). The Importance of on-going concurrent activities as a function of age in time- and event-based prospective memory. *European Journal of Cognitive Psychology, 11*, 219-237. doi: 10.1080/713752309

de Jager, C. A., Budge, M. M., & Clarke, R. (2003). Utility of TICS-M for the assessment of cognitive function in older adults. *International Journal of Geriatric Psychiatry, 18*, 318-324. doi: 10.1002/gps.830

Deary, I. J., Liewald, D., & Nissan, J. (2011). A free, easy-to-use, computer-based simple and four-choice reaction time programme: The Deary-Liewald reaction time task. *Behavior Research Methods, 43*, 258-268. doi: 10.3758/s13428-010-0024-1

Dobbs, A. R., & Rule, B. G. (1987). Prospective memory and self-reports of memory abilities in older adults. *Canadian Journal of Psychology, 41*, 209. doi: 10.1037/h0084152

Duchek, J. M., Balota, D. A., & Cortese, M. (2006). Prospective memory and apolipoprotein E in healthy aging and early stage Alzheimer's disease. *Neuropsychology, 20*, 633-644. doi: 10.1037/0894-4105.20.6.633

Ebbinghaus, H. (1885/2013). Memory: A contribution to experimental psychology. *Annals of Neurosciences, 20*, 155-156. doi: 10.5214/ans.0972.7531.200408

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. doi: 10.1037/0882-7974.7.3.471

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 717. doi: 10.1037//0278-7393.16.4.717

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science, 14,* 286-290. doi: 10.1111/j.0963-7214.2005.00382.x

Einstein, G. O., McDaniel, M. A., Manzi, M., Cochran, B., & Baker, M. (2000). Prospective memory and aging: Forgetting intentions over short delays. *Psychology and Aging, 15*, 671-683. doi: 10.1037/0882-7974.15.4.671

Einstein, G. O., McDaniel, M. A., Richardson, S. L., Guynn, M. J., & Cunfer, A. R. (1995). Aging and prospective memory: Examining the influences of self-initiated retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 996-1007. doi: 10.1037/0278-7393.21.4.996

Ellis, J. A. (1996). Prospective memory or the realization of delayed intentions: A conceptual framework for research. In M. A. Brandimonte, G. O. Einstein & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 1-22). Mahwah, NJ: Lawrence Erlbaum.

Ellis, J. A., & Kvavilashvili, L. (2000). Prospective memory in 2000: Past, present, and future directions. *Applied Cognitive Psychology, 14*, S1-S9. doi: 10.1002/acp.767

Farina, N., Young, J., Tabet, N., & Rusted, J. (2013). Prospective memory in Alzheimer-type dementia: Exploring prospective memory performance in an age-stratified sample. *Journal of Clinical and Experimental Neuropsychology, 35*, 1-10. doi: 10.1080/13803395.2013.844772

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189-198. doi: 10.1016/0022-3956(75)90026-6

Foster, E. R., McDaniel, M. A., & Rendell, P. G. (2017). Improving prospective memory in persons with Parkinson Disease: A randomized controlled trial. *Neurorehabilitation and Neural Repair, 31*, 451-461. doi: 10.1177/1545968317690832

Foster, E. R., Rose, N. S., McDaniel, M. A., & Rendell, P. G. (2013). Prospective memory in Parkinson disease during a virtual week: Effects of both prospective and retrospective demands. *Neuropsychology, 27,* 170–194. doi: 10.1037/a0031946.

Freud, S. (1901/1952). *The psychopathology of everyday life*. London: Hogarth.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133*, 101-135. doi: 10.1037/0096-3445.133.1.101

Gardner, R. S., & Ascoli, G. A. (2015). The natural frequency of human prospective memory increases with age. *Psychology and Aging, 30,* 209-219. doi: 10.1037/a0038876 / 10.1037/a0038876.supp (Supplemental)

Gilbert, S. J. (2015). Strategic offloading of delayed intentions into the external environment. *The Quarterly Journal of Experimental Psychology, 68,* 971-992. doi: 10.1080/17470218.2014.972963

Gonneaud, J., Kalpouzos, G., Bon, L., Viader, F., Eustache, F., & Desgranges, B. (2011). Distinct and shared cognitive functions mediate event-and time-based prospective memory impairment in normal ageing. *Memory, 19*, 360-377. doi: 10.1080/09658211.2011.570765

Groot, Y. C. T., Wilson, B. A., Evans, J., & Watson, P. (2002). Prospective memory functioning in people with and without brain injury. *Journal of International Neuropsychology Society, 8*, 645-654. doi: 10.1017/S1355617702801321

Habota, T., McLennan, S. N., Cameron, J., Henry, J. D., Ski, C. F., Thompson, D. R., & Rendell, P. G. (2015). Prospective memory impairment in chronic heart failure. *Journal of the International Neuropsychological Society, 21*, 183-192. doi: 10.1017/S1355617715000119

Haines, S. J., Shelton, J. T., Henry, J. D., Terrett, G., Vorwerk, T., & Rendell, P. G. (2019). Prospective memory and cognitive aging *Oxford Research Encyclopedia of Psychology*: Oxford University Press. doi: 10.1093/acrefore/9780190236557.013.381

Harris, J. E., & Wilkins, A. J. (1982). Remembering to do things: A theoretical framework and an illustrative experiment. *Human Learning, 1*, 123-136.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*, 356-388. doi: 10.1037/0096-3445.108.3.356

Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1990). Aging and qualitative characteristics of memories for perceived and imagined complex events. *Psychology and Aging, 5*, 119. doi: 10.1037/0882-7974.5.1.119

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.): New York: Guilford Press.

Heathcote, A., Loft, S., & Remington, R. W. (2015). Slow Down and Remember to Remember! A Delay Theory of Prospective Memory Costs. *Psychological Review, 122*, 376-410. doi: 10.1037/a0038952

Henry, J. D., MacLeod, M. S., Phillips, L. H., & Crawford, J. R. (2004). A meta-analytic review of prospective memory and aging. *Psychology and Aging, 19*, 27-39. doi: 10.1037/0882-7974.19.1.27

Henry, J. D., & Phillips, L. H. (2006). Covariates of production and perseveration on tests of phonemic, semantic and alternating fluency in normal aging. *Aging, Neuropsychology, and Cognition, 13*, 529-551. doi: 10.1080/138255890969537

Henry, J. D., Rendell, P. G., Kliegel, M., & Altgassen, M. (2007). Prospective memory in schizophrenia: Primary or secondary impairment? *Schizophrenia Research, 95*, 179-185. doi: 10.1016/j.schres.2007.06.003

Henry, J. D., Rendell, P. G., Phillips, L. H., Dunlop, L., & Kliegel, M. (2012). Prospective memory reminders: A laboratory investigation of initiation source and age effects. *The Quarterly Journal of Experimental Psychology, 65*, 1274-1287. doi: 10.1080/17470218.2011.651091

Henry, J. D., Rendell, P. G., Rogers, P., Altgassen, M., & Kliegel, M. (2012). Prospective memory in schizophrenia and schizotypy. *Cognitive Neuropsychiatry, 17*, 133-150. doi: 10.1080/13546805.2011.581536

Hering, A., Cortez, S. A., Kliegel, M., & Altgassen, M. (2014). Revisiting the age-prospective memory-paradox: The role of planning and task experience. *European Journal of Ageing, 11*, 99-106. doi: 10.1007/s10433-013-0284-6

Hering, A., Kliegel, M., Rendell, P. G., Craik, F. I. M., & Rose, N. S. (2018). Prospective memory is a key predictor of functional independence in older adults. *Journal of the International Neuropsychological Society, 24*, 640-645. doi: 10.1017/S1355617718000152

Hertzog, C., Dixon, R. A., Hultsch, D. F., & Macdonald, S. W. S. (2003). Latent change models of adult cognition: Are changes in processing speed and working memory associated with changes in episodic memory? *Psychology and Aging, 18*, 755-769. doi: 10.1037/0882-7974.18.4.755

Hertzog, C., & Dunlosky, J. (2011). Metacognition in Later Adulthood: Spared Monitoring Can Benefit Older Adults' Self-Regulation. *Current Directions in Psychological Science, 20*, 167-173. doi: 10.1177/0963721411409026

Hofer, S. M., & Alwin, D. F. (2008). *Handbook of cognitive aging: Interdisciplinary perspectives*. Thousand Oaks, CA: Sage.

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2014). Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity. *Behavior Research Methods, 46*, 702-721. doi: 10.3758/s13428-013-0411-5

Huppert, F. A., Johnson, T., & Nickson, J. (2000). High prevalence of prospective memory impairment in the elderly and in early-stage dementia: Findings from a population-based study. *Applied Cognitive Psychology, 14*, S63-S81. doi: 10.1002/acp.771

Ihle, A., Ghisletta, P., & Kliegel, M. (2016). Prospective memory and intraindividual variability in ongoing task response times in an adult lifespan sample: The role of cue focality. *Memory*, 1-7. doi: 10.1080/09658211.2016.1173705

Ihle, A., Schnitzspahn, K., Rendell, P. G., Luong, C., & Kliegel, M. (2012). Age benefits in everyday prospective memory: The influence of personal task importance, use of reminders and everyday stress. *Aging, Neuropsychology, and Cognition, 19*, 84-101. doi: 10.1080/13825585.2011.629288

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory, 18*, 394-412. doi: 10.1080/09658211003702171

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3-28. doi: 10.1037/0033-2909.114.1.3

Joly-Burra, E., Van der Linden, M., & Ghisletta, P. (2018). Intraindividual variability in inhibition and prospective memory in healthy older adults: Insights from response regularity and rapidity. *Journal of Intelligence, 6*, 13. doi: 10.3390/jintelligence6010013

Jones, S., Livner, Å., & Bäckman, L. (2006). Patterns of prospective and retrospective memory impairment in preclinical Alzheimer's disease. *Neuropsychology, 20*, 144-152. doi: 10.1037/0894-4105.20.2.144

Kamat, R., Weinborn, M., Kellogg, E. J., Bucks, R. S., Velnoweth, A., & Woods, S. P. (2014). Construct validity of the Memory for Intentions Screening Test (MIST) in healthy older adults. *Assessment, 21*, 742-753. doi: 10.1177/1073191114530774

Kelemen, W. L., Weinberg, W. B., Alford, H. S., Mulvey, E. K., & Kaeochinda, K. F. (2006). Improving the reliability of event-based laboratory tests of prospective memory. *Psychonomic Bulletin and Review, 13*, 1028-1032. doi: 10.3758/bf03213920

Kidder, D. P., Park, D. C., Hertzog, C., & Morrell, R. W. (1997). Prospective memory and aging: The effects of working memory and prospective memory task load. *Aging, Neuropsychology, and Cognition, 4*, 93-112. doi: 10.1080/13825589708256639

Kliegel, M., Ballhausen, N., Hering, A., Ihle, A., Schnitzspahn, K., & Zuber, S. (2016). Prospective memory in older adults: Where we are now and what is next. *Gerontology, 62*, 459-466. doi: 10.1159/000443698

Kliegel, M., & Jager, T. (2006). Delayed–execute prospective memory performance: The effects of age and working memory. *Developmental Neuropsychology, 30*, 819-843. doi: 10.1207/s15326942dn3003_4

Kliegel, M., Jäger, T., & Phillips, L. H. (2008). Adult age differences in event-based prospective memory: A meta-analysis on the role of focal versus nonfocal cues. *Psychology and Aging, 23*, 203–208. doi: 10.1037/0882-7974.23.1.203

Kliegel, M., MacKinlay, R., & Jäger, T. (2008). A life span approach to the development of complex prospective memory. In M. Kliegel, M. A. McDaniel & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 187-216). New York: Lawrence Erlbaum.

Kliegel, M., & Martin, M. (2003). Prospective memory research: Why is it relevant? *International Journal of Psychology, 38*, 193-194. doi: 10.1080/00207590244000205

Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2002). Complex prospective memory and executive control of working memory: A process model. *Psychological Test and Assessment Modeling, 44*, 303–318.

Kliegel, M., Martin, M., McDaniel, M. A., & Phillips, L. H. (2007). Adult age differences in errand planning: The role of task familiarity and cognitive resources. *Experimental Aging Research, 33*, 145-161. doi: 10.1080/03610730601177395

Kliegel, M., McDaniel, M. A., & Einstein, G. O. (2000). Plan formation, retention, and execution in prospective memory: A new approach and age-related effects. *Memory and Cognition, 28*, 1041-1049. doi: 10.3758/bf03209352

Kliegel, M., McDaniel, M. A., & Einstein, G. O. (2008). *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives*. New York: Lawrence Erlbaum.

Kliegel, M., Rendell, P. G., & Altgassen, M. (2008). The added value of an applied perspective in cognitive gerontology. In S. M. Hofer & D. F. Alwin (Eds.), *Handbook of cognitive aging: Interdisciplinary perspectives* (pp. 587-602). Thousand Oaks, CA: Sage.

Koriat, A., Ben-Zur, H., & Sheffer, D. (1988). Telling the same story twice: Output monitoring and age. *Journal of Memory and Language, 27*, 23-39. doi: 10.1016/0749-596X(88)90046-0

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging, 15*, 126-147. doi: 10.1037/0882-7974.15.1.126

Kvavilashvili, L. (1987). Remembering intention as a distinct form of memory. *British Journal of Psychology, 78*, 507-518. doi: 10.1111/j.2044-8295.1987.tb02265.x

Kvavilashvili, L., Cockburn, J., & Kornbrot, D. E. (2013). Prospective memory and ageing paradox with event-based tasks: A study of young, young-old, and old-old participants. *The Quarterly Journal of Experimental Psychology, 66*, 864-875. doi: 10.1080/17470218.2012.721379

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. doi: 10.1037/0096-3445.136.1.112

Lacruz, M., Emeny, R., Bickel, H., Linkohr, B., & Ladwig, K. (2013). Feasibility, internal consistency and covariates of TICS-m (Telephone Interview for Cognitive Status-modified) in a population-based sample: Findings from the KORA-age study. *International Journal of Geriatric Psychiatry, 28*, 971-978. doi: 10.1002/gps.3916

Langenecker, S. A., Zubieta, J.-K., Young, E. A., Akil, H., & Nielson, K. A. (2007). A task to manipulate attentional load, set-shifting, and inhibitory control: Convergent validity and test–retest reliability of the Parametric Go/No-Go Test. *Journal of Clinical and Experimental Neuropsychology, 29*, 842-853. doi: 10.1080/13803390601147611

Lee, J. H., Shelton, J. T., Scullin, M. K., & McDaniel, M. A. (2016). An implementation intention strategy can improve prospective memory in older adults with very mild Alzheimer's disease. *British Journal of Clinical Psychology, 55*, 154-166. doi: 10.1111/bjc.12084

Lee, S. D., Ong, B., Pike, K. E., & Kinsella, G. J. (2018). Prospective memory and subjective memory decline: A neuropsychological indicator of memory difficulties in community-dwelling older people. *Journal of Clinical and Experimental Neuropsychology, 40*, 183-197. doi: 10.1080/13803395.2017.1326465

Leitz, J. R., Morgan, C. J., Bisby, J. A., Rendell, P. G., & Curran, H. V. (2009). Global impairment of prospective memory following acute alcohol. *Psychopharmacology, 205*, 379-387. doi: 10.1007/s00213-009-1546-z

Lewin, K. (1926/1999). Intention, will, and need. In M. Gold (Ed.), *The complete social scientist: A Kurt Lewin reader.* (pp. 83-115). Washington, DC: American Psychological Association.

Logie, R., Maylor, E., Della Sala, S., & Smith, G. (2004). Working memory in event- and time-based prospective memory tasks: Effects of secondary demand and age. *European Journal of Cognitive Psychology, 16*, 441-456. doi: 10.1080/09541440340000114

Maillet, D., & Schacter, D. L. (2016). From mind wandering to involuntary retrieval: Age-related differences in spontaneous cognitive processes. *Neuropsychologia*, *80*, 142-156. doi: 10.1016/j.neuropsychologia.2015.11.017

Marsh, R. L., & Hicks, J. L. (1998). Event-based prospective memory and executive control of working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 336-349. doi: 10.1037/0278-7393.24.2.336

Marsh, R. L., & Hicks, J. L. (2002). Comparisons of target output monitoring and source input monitoring. *Applied Cognitive Psychology, 16*, 845-862. doi: 10.1002/acp.859

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2008). On beginning to understand the role of context in prospective memory. In M. Kliegel, M. A. McDaniel & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 77-100). New York: Lawrence Erlbaum.

Marsh, R. L., Hicks, J. L., Cook, G. I., & Mayhorn, C. B. (2007). Comparing Older and Younger Adults in an Event-Based Prospective Memory Paradigm Containing an Output Monitoring Component. *Aging, Neuropsychology, and Cognition, 14*, 168-188. doi: 10.1080/138255891007074

Marsh, R. L., Hicks, J. L., Cook, G. I., Hansen, J. S., & Pallos, A. L. (2003). Interference to ongoing activities covaries with the characteristics of an event-based intention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 861-870. doi: 10.1037/0278-7393.29.5.861

Marsh, R. L., Hicks, J. L., & Hancock, T. W. (2000). On the interaction of ongoing cognitive activity and the nature of an event-based intention. *Applied Cognitive Psychology*, *14*, S29-S41. doi: 10.1002/acp.769

Marsh, R. L., Hicks, J. L., & Landau, J. D. (1998). An investigation of everyday prospective memory. *Memory and Cognition, 26*, 633-643. doi: 10.3758/BF03211383

Martin, M., Kliegel, M., & McDaniel, M. A. (2003). The involvement of executive functions in prospective memory performance of adults. *International Journal of Psychology, 38*, 195-206. doi: 10.1080/00207590244000205

Masumoto, K., Nishimura, C., Tabuchi, M., & Fujita, A. (2011). What factors influence prospective memory for elderly people in a naturalistic setting? *Japanese Psychological Research, 53*, 30-41. doi: 10.1111/j.1468-5884.2010.00453.x

Maylor, E. A. (1990). Age and prospective memory. *The Quarterly Journal of Experimental Psychology Section A, 42*, 471-493. doi: 10.1080/14640749008401233

Maylor, E. A. (1993). Aging and forgetting in prospective and retrospective memory tasks. *Psychology and Aging, 8*, 420. doi: 10.1037/0882-7974.8.3.420

Maylor, E. A. (1996). Age-related impairment in an event-based prospective-memory task. *Psychology and Aging, 11*, 74–78. doi: 10.1037/0882-7974.11.1.74

Maylor, E. A., Smith, G., Della Sala, S., & Logie, R. H. (2002). Prospective and retrospective memory in normal aging and dementia: An experimental study. *Memory and Cognition, 30*, 871-884. doi: 10.3758/BF03195773

McAlister, C., & Schmitter-Edgecombe, M. (2013). Naturalistic assessment of executive function and everyday multitasking in healthy older adults. *Aging, Neuropsychology, and Cognition, 20*, 735-756. doi: 10.1080/13825585.2013.781990

McDaniel, M. A., Binder, E. F., Bugg, J. M., Waldum, E. R., Dufault, C., Meyer, A., . . . Kudelka, C. (2014). Effects of cognitive training with and without aerobic exercise on cognitively demanding everyday activities. *Psychology and Aging, 29,* 717. doi: 10.1037/a0037363

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology, 14,* S127-S144. doi: https://doi-org.ezproxy1.acu.edu.au/10.1002/acp.775

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field.* Thousand Oaks, CA: Sage.

McDaniel, M. A., & Einstein, G. O. (2008). Prospective memory and aging: Old issues and new questions. In S. M. Hofer & D. F. Alwin (Eds.), *Handbook of cognitive aging: Interdisciplinary perspectives*. Thousand Oaks, CA: Sage.

McDaniel, M. A., Einstein, G. O., Stout, A. C., & Morgan, Z. (2003). Aging and maintaining intentions over delays: do it or lose it. *Psychology and Aging, 18,* 823. doi: 10.1037/0882-7974.18.4.823

McDaniel, M. A., Glisky, E. L., Guynn, M. J., & Routhieaux, B. C. (1999). Prospective memory: A neuropsychological study. *Neuropsychology, 13*, 103. doi: 10.1037/0894-4105.13.1.103

McDaniel, M. A., Shelton, J. T., Breneiser, J. E., Moynan, S., & Balota, D. A. (2011). Focal and nonfocal prospective memory performance in very mild dementia: A signature decline. *Neuropsychology, 25*, 387. doi: 10.1037/a0021682

McDaniel, M. A., Umanath, S., Einstein, G. O., & Waldum, E. R. (2015). Dual pathways to prospective remembering. *Frontiers in Human Neuroscience, 9*. doi: 10.3389/fnhum.2015.00392

Meacham, J. A., & Singer, J. (1977). Incentive effects in prospective remembering. *The Journal of Psychology, 97*, 191-197. doi: 10.1080/00223980.1977.9923962

Miller, K. M., Price, C. C., Okun, M. S., Montijo, H., & Bowers, D. (2009). Is the n-back task a valid neuropsychological measure for assessing working memory? *Archives of Clinical Neuropsychology*, *24*, 711-717. doi: 10.1093/arclin/acp063

Mioni, G., Bertucci, E., Rosato, A., Terrett, G., Rendell, P. G., Zamuner, M., & Stablum, F. (2017). Improving prospective memory performance with future event simulation in traumatic brain injury patients. *British Journal of Clinical Psychology*, *2*, 130-148. doi: 10.1111/bjc.12126

Mioni, G., Cardullo, S., Ciavarelli, A., & Stablum, F. (2019). Age-related changes in time discrimination: The involvement of inhibition, working memory and speed of processing. *Current Psychology*, 1-10. doi: 10.1007/s12144-019-00170-8

Mioni, G., Rendell, P. G., Stablum, F., Gamberini, L., & Bisiacchi, P. S. (2015). Test–retest consistency of Virtual Week: A task to investigate prospective memory. *Neuropsychological Rehabilitation*, *25*, 419-447. doi: 10.1080/09602011.2014.941295

Mioni, G., & Stablum, F. (2014). Monitoring behaviour in a time-based prospective memory task: The involvement of executive functions and time perception. *Memory, 22*, 536-552. doi: 10.1080/09658211.2013.801987

Mioni, G., Stablum, F., Biernacki, K., & Rendell, P. G. (2015). Virtual Week: Translation and adaptation for the Italian population. *Neuropsychological Rehabilitation*, *27*, 1-21. doi: 10.1080/09602011.2015.1103758

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100. doi: 10.1006/cogp.1999.0734

Moscovitch, M. (1982). A neuropsychological approach to perception and memory in normal and pathological aging. In F. I. M. Craik & S. Trehub (Eds.), *Aging and Cognitive Processes* (pp. 55-78). New York: Plenum Press.

Mullet, H. G., Scullin, M. K., Hess, T. J., Scullin, R. B., Arnold, K. M., & Einstein, G. O. (2013). Prospective memory and aging: Evidence for preserved spontaneous retrieval with exact but not related cues. *Psychology and Aging, 28*, 910. doi: 10.1037/a0034347

Nelson, H. E. (1982). *National Adult Reading Test (NART): For the assessment of premorbid intelligence in patients with dementia: Test manual*: NFER-Nelson.

Nelson, H. E., & Willison, J. (1991). *The National Adult Reading Test (NART): Updated manual*: The NFER-Nelson.

Niedźwieńska, A., & Barzykowski, K. (2012). The age prospective memory paradox within the same sample in time-based and event-based tasks. *Aging, Neuropsychology, and Cognition, 19*, 58-83. doi: 10.1080/13825585.2011.628374

Niedźwieńska, A., Janik, B., & Jarczyńska, A. (2013). Age-related differences in everyday prospective memory tasks: The role of planning and personal importance. *International Journal of Psychology, 48*, 1291-1302. doi: 10.1080/00207594.2012.752097

Niedźwieńska, A., Kvavilashvili, L., Ashaye, K., & Neckar, J. (2017). Spontaneous retrieval deficits in amnestic mild cognitive impairment: A case of focal event-based prospective memory. *Neuropsychology, 31*, 735-749. doi: 10.1037/neu0000378

Niedźwieńska, A., Rendell, P. G., Barzykowski, K., & Leszczyńska, A. (2014). Only social feedback reduces age-related prospective memory deficits in “Virtual Week”. *International Psychogeriatrics, 26,* 759-767. doi: 10.1017/S1041610214000027

Niedźwieńska, A., Rendell, P. G., Barzykowski, K., & Leszczyńska, A. (2016). Virtual Week: Validity and psychometric properties of a Polish adaptation. *European Review of Applied Psychology, 66,* 79-84. doi: 10.1016

Ozgis, S., Rendell, P. G., & Henry, J. D. (2009). Spaced retrieval significantly improves prospective memory performance of cognitively impaired older adults. *Gerontology, 55*, 229-232. doi: 10.1159/000163446

Pankratz, V. S., Roberts, R. O., Mielke, M. M., Knopman, D. S., Jack, C. R., Geda, Y. E., . . . Petersen, R. C. (2015). Predicting the risk of mild cognitive impairment in the Mayo clinic study of aging. *Neurology, 84*, 1433-1442. doi: 10.1212/WNL.0000000000001437

Park, D. C., Hertzog, C., Kidder, D. P., & Morrell, R. W. (1997). Effect of age on event-based and time-based prospective memory. *Psychology and Aging, 12,* 314–327. doi: 10.1037/0882-7974.12.2.314

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging, 17,* 299-320. doi: 10.1037/0882-7974.17.2.299

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology, 60*, 173-196. doi: 10.1146/annurev.psych.59.103006.093656

Patton, G. W., & Meit, M. (1993). Effect of aging on prospective and incidental memory. *Experimental Aging Research, 19,* 165-176. doi: 10.1080/03610739308253929

Pereira, A., Ellis, J. A., & Freeman, J. E. (2012). The effects of age, enactment, and cue-action relatedness on memory for intentions in the Virtual Week task. *Aging, Neuropsychology, and Cognition, 19*, 1-17. doi: 10.1080/13825585.2011.638977

Phillips, L. H., Bull, R., Allen, R., Insch, P., Burr, K., & Ogg, W. (2011). Lifespan aging and belief reasoning: Influences of executive function and social cue decoding. *Cognition, 120*, 236-247. doi: 10.1016/j.cognition.2011.05.003

Phillips, L. H., & Henry, J. D. (2010). Adult aging and executive functioning. In V. Anderson, R. Jacobs & P. J. Anderson (Eds.), *Executive functions and the frontal lobes: A lifespan perspective*. New York: Taylor & Francis.

Phillips, L. H., Henry, J. D., & Martin, M. (2008). Adult aging and prospective memory: The importance of ecological validity. In M. Kliegel, M. A. McDaniel & G. O. Einstein (Eds.), (pp. 161-181). New York: Lawrence Erlbaum.

Rabbitt, P. (1996). Commentary: Why are studies of 'prospective memory' planless? In M. Brandimonte, G. O. Einstein & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 239-248). Mahwah, NJ: Lawrence Erlbaum.

Randall, S. E. (2016). *Age differences in prospective memory: Laboratory versus naturalistic settings.* (Doctoral thesis), Australian Catholic University, Melbourne, Australia.

Raskin, S., Buckheit, C., & Sherrod, C. (2010). *Memory for Intentions Test*. Lutz, FL: PAR.

Reichman, W. E., Fiocco, A. J., & Rose, N. S. (2010). Exercising the brain to avoid cognitive decline: Examining the evidence. *Aging Health, 6*, 565-584. doi: 10.2217/AHE.10.54

Rendell, P. G., & Craik, F. I. (2000). Virtual week and actual week: Age-related differences in prospective memory. *Applied Cognitive Psychology, 14*, 43-62. doi: 10.1002/acp.770

Rendell, P. G., & Henry, J. D. (2009). A review of Virtual Week for prospective memory assessment: Clinical implications. *Brain Impairment*, *10*, 14-22. doi: 10.1375/brim.10.1.14

Rendell, P. G., Jensen, F., & Henry, J. D. (2007). Prospective memory in multiple sclerosis. *Journal of the International Neuropsychological Society*, *13*, 410-416. doi: 10.10170S1355617707070579

Rendell, P. G., McDaniel, M. A., Forbes, R. D., & Einstein, G. O. (2007). Age-related effects in prospective memory are modulated by ongoing task complexity and relation to target cue. *Aging, Neuropsychology, and Cognition*, *14*, 236-256. doi: 10.1080/13825580600579186

Rendell, P. G., Phillips, L. H., Henry, J. D., Brumby-Rendell, T., de la Piedad Garcia, X., Altgassen, M., & Kliegel, M. (2011). Prospective memory, emotional valence and ageing. *Cognition and Emotion*, *25*, 916-925. doi: 10.1080/02699931.2010.508610

Rendell, P. G., & Thomson, D. M. (1993). The effect of ageing on remembering to remember: An investigation of simulated medication regimens. *Australian Journal on Ageing*, *12*, 11-18. doi: 10.1111/j.1741-6612.1993.tb00578.x

Rendell, P. G., & Thomson, D. M. (1999). Aging and prospective memory: Differences between naturalistic and laboratory tasks. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *54*, 256-269. doi: 10.1093/geronb/54B.4.P256

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20*, 676-688. doi: 10.1016/j.tics.2016.07.002

Roeckelein, J., & Roeckelein, J. E. (2006). *Elsevier's Dictionary of Psychological Theories*. Amsterdam: Elsevier Science & Technology.

Rose, N. S., Myerson, J., Sommers, M. S., & Hale, S. (2009). Are there age differences in the executive component of working memory? Evidence from domain-general interference effects. *Aging, Neuropsychology, and Cognition, 16*, 633-653. doi: 10.1080/13825580902825238

Rose, N. S., Rendell, P. G., Hering, A., Kliegel, M., Bidelman, G. M., & Craik, F. I. (2015). Cognitive and neural plasticity in older adults' prospective memory following training with the Virtual Week computer game. *Frontiers in Human Neuroscience, 9*. doi: 10.3389/fnhum.2015.00592

Rose, N. S., Rendell, P. G., McDaniel, M. A., Aberle, I., & Kliegel, M. (2010). Age and individual differences in prospective memory during a "Virtual Week": The roles of working memory, vigilance, task regularity, and cue focality. *Psychology and Aging, 25*, 595. doi: 10.1037/a0019771

Salthouse, T. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science, 2*, 179-183. doi: 10.1111/j.1467-9280.1991.tb00127.x

Salthouse, T. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403-428. doi: 10.1037/0033-295X.103.3.403

Salthouse, T. (2000). Aging and measures of processing speed. *Biological Psychology, 54*, 35-54. doi: 10.1016/S0301-0511(00)00052-1

Salthouse, T. (2014). Correlates of cognitive change. *Journal of Experimental Psychology: General, 143*, 1026. doi: 10.1037/a0034847

Salthouse, T. (2016). Continuity of cognitive change across adulthood. *Psychonomic Bulletin Review, 23*, 932-939. doi: 10.3758/s13423-015-0910-8

Salthouse, T., Berish, D. E., & Siedlecki, K. L. (2004). Construct validity and age sensitivity of prospective memory. *Memory and Cognition*, *32*, 1133-1148. doi: 10.3758/BF03196887

Schaie, K. W. (1996). *Intellectual development in adulthood: The Seattle longitudinal study*: Cambridge University Press.

Schmidt, W., Berg, I., Betto, G., & Deelman, I. (2001). Prospective memory training in older adults. *Educational Gerontology, 27,* 455-478. doi: 10.1080/036012701316894162

Schnitzspahn, K. M., Ihle, A., Henry, J. D., Rendell, P. G., & Kliegel, M. (2011). The age-prospective memory-paradox: An exploration of possible mechanisms. *International Psychogeriatrics*, *23*, 583-592. doi: 10.1017/S1041610210001651

Schnitzspahn, K. M., Stahl, C., Zeintl, M., Kaller, C. P., & Kliegel, M. (2013). The role of shifting, updating, and inhibition in prospective memory performance in young and older adults. *Developmental Psychology*, *49*, 1544–1553. doi: 10.1037/a0030579

Schryer, E., & Ross, M. (2013). The use and benefits of external memory aids in older and younger adults. *Applied Cognitive Psychology, 27*, 663-671. doi: 10.1002/acp.2946

Scullin, M. K., McDaniel, M. A., & Shelton, J. T. (2013). The dynamic multiprocess framework: Evidence from prospective memory with contextual variability. *Cognitive Psychology*, *67*, 55-71. doi: 10.1016/j.cogpsych.2013.07.001.

Sellen, A. J., Louie, G., Harris, J., & Wilkins, A. (1997). What brings intentions to mind? An in situ study of prospective memory. *Memory, 5*, 483-507.

Shelton, J. T., & Christopher, E. A. (2016). A fresh pair of eyes on prospective memory monitoring. *Memory and Cognition, 44,* 837-845. doi: 10.3758/s13421-016-0601-3

Shelton, J. T., Lee, J. H., Scullin, M. K., Rose, N. S., Rendell, P. G., & McDaniel, M. A. (2016). Improving prospective memory in healthy older adults and individuals with very mild Alzheimer's disease. *Journal of the American Geriatrics Society, 64*, 1307-1312. doi: 10.1111/jgs.14134

Shelton, J. T., & Scullin, M. K. (2017). The dynamic interplay between bottom-up and top-down processes supporting prospective remembering. *Current Directions in Psychological Science, 26*, 352-358. doi: 10.1177/0963721417700504

Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology, 33*, 5-42. doi: 10.1006/cogp.1997.0658

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 347-361. doi: 10.1037/0278-7393.29.3.347

Smith, R. E. (2008). Connecting the past and the future: Attention, memory and delayed intentions. In M. Kliegel, M. A. McDaniel & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives*. New York: Lawrence Erlbaum.

Smith, R. E., & Bayen, U. J. (2004). A multinomial model of event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 756. doi: 10.1037/0278-7393.30.4.756

Smith, R. E., Horn, S. S., & Bayen, U. J. (2012). Prospective memory in young and older adults: The effects of ongoing-task load. *Aging, Neuropsychology, and Cognition, 19*, 495-514. doi: 10.1080/13825585.2011.633161.

Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *The Lancet Neurology, 11*, 1006-1012. doi: 10.1016/S1474-4422(12)70191-6

Strickland, L., Loft, S., Remington, R. W., & Heathcote, A. (2018). Racing to remember: A theory of decision control in event-based prospective memory. *Psychological Review, 125*, 851-887. doi: 10.1037/rev0000113

Tam, J. W., & Schmitter-Edgecombe, M. (2013). Event-based prospective memory and everyday forgetting in healthy older adults and individuals with mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, *35*, 279-290. doi: 10.1080/13803395.2013.770823

Terrett, G., Horner, K., White, R., Henry, J. D., Kliegel, M., Labuschagne, I., & Rendell, P. G. (2019). The relationship between episodic future thinking and prospective memory in middle childhood: Mechanisms depend on task type. *Journal of Experimental Child Psychology, 178*, 198-213. doi: 10.1016/j.jecp.2018.10.003

Terrett, G., Rose, N. S., Henry, J. D., Bailey, P. E., Altgassen, M., Phillips, L. H., . . . Rendell, P. G. (2015). The relationship between prospective memory and episodic future thinking in younger and older adulthood. *The Quarterly Journal of Experimental Psychology*, 1-34. doi: 10.1080/17470218.2015.1054294

Thompson, C., Henry, J. D., Rendell, P. G., Withall, A., & Brodaty, H. (2010). Prospective memory function in mild cognitive impairment and early dementia. *Journal of the International Neuropsychological Society, 16*, 318-325. doi: 10.1017/S1355617709991354

Tse, C., Chang, J., Fung, A., Lam, L., Hau, K., Leung, G., & Balota, D. (2015). The utility of a non-verbal prospective memory measure as a sensitive marker for early-stage Alzheimer's disease in Hong Kong. *International Psychogeriatrics, 27*, 231-242. doi: 10.1017/S1041610214002038

Uttl, B. (2008). Transparent meta-analysis of prospective memory and aging. *PloS one, 3*, e1568. doi: 10.1371/journal.pone.0001568

Verhaeghen, P. (2014). *The elements of cognitive aging: Meta-analyses of age-related differences in processing speed and their consequences*: New York: Oxford University Press.

Volle, E., Gonen-Yaacovi, G., de Lacy Costello, A., Gilbert, S. J., & Burgess, P. W. (2011). The role of rostral prefrontal cortex in prospective memory: A voxel-based lesion study. *Neuropsychologia, 49*, 2185-2198. doi: 10.1016/j.neuropsychologia.2011.02.045

Waldum, E. R., Dufault, C. L., & McDaniel, M. A. (2016). Prospective memory training. *Journal of Applied Gerontology, 35*, 1211-1234. doi: 10.1177/0733464814559418

West, R. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin, 120*, 272-292. doi: 10.1037/0033-2909.120.2.272

West, R., & Craik, F. I. (2001). Influences on the efficiency of prospective memory in younger and older adults. *Psychology and Aging, 16*, 682-696. doi: 10.1037/0882-7974.16.4.682

Will, C. M., Rendell, P. G., Ozgis, S., Pierson, J. M., Ong, B., & Henry, J. D. (2009). Cognitively impaired older adults exhibit comparable difficulties on naturalistic and laboratory prospective memory tasks. *Applied Cognitive Psychology, 23,* 804-812. doi: 10.1002/acp.1514

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (1996). *Behavioural Assessment of the Dysexecutive Syndrome*. London: Pearson Assessment.

Wilson, B. A., Emslie, H., Foley, J., Shiel, A., Watson, P., Hawkins, K., . . . Evans, J. J. (2005). *The Cambridge Prospective Memory Test*. London: Pearson.

Zeintl, M., Kliegel, M., & Hofer, S. M. (2007). The role of processing resources in age-related prospective and retrospective memory within old age. *Psychology and Aging, 22*, 826-834. doi: 10.1037/0882-7974.22.4.826 / 10.1037/0882-7974.22.4.826.supp (Supplemental)

Zuber, S., Kliegel, M., & Ihle, A. (2016). An individual difference perspective on focal versus nonfocal prospective memory. *Memory and Cognition, 44*, 1192-1203. doi: 10.3758/s13421-016-0628-5