

**Brain-based versus External Memory Stores:
Influencing Factors and Underlying Neural Correlates**

Pei-Chun Tsai

Supervisor: Prof. Sam J. Gilbert

Institute of Cognitive Neuroscience

Faculty of Brain Sciences

University College London (UCL)

Thesis submitted for the degree of

Doctor of Philosophy in Cognitive Neuroscience

November 2022

Declaration:

I, Pei-Chun Tsai, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Date: 2022/11/18

Abstract

Technological advancements provide people with more opportunity to rely on external resources to support cognitive processes. These associated processes are defined as cognitive offloading (Risko & Gilbert, 2016). The current thesis aims to explore the psychological processes and neural mechanism of cognitive offloading. In Experiment 1, we developed an 'optimal reminder' task by calculating whether people were biased towards using reminders or their own memory, compared with an optimal strategy. If participants were biased, the second purpose of Experiment 1 was to assess whether such bias could be reduced through metacognitive advice. Results revealed people were biased towards setting reminders, and the bias was eliminated by metacognitive advice. Experiment 2 used the optimal reminder task to evaluate the effect of ageing on cognitive offloading. This showed that older people set more reminders than younger adults, but were less biased towards setting reminders when the impaired memory performance of older people was taken into account. Experiment 3 investigated the effects of three factors: delay length, metacognitive judgement, and clock revealability, on cognitive offloading in a time-based task (e.g. remembering to press a specific button after 10 seconds). We found participants' use of reminders was based on both the characteristics of the task (i.e., delay and clock revealability) and metacognitive judgements. Experiment 4 used fMRI to evaluate whether an instruction to offload information to an external reminder triggered different brain activity to an instruction to forget or remember. Results showed that brain activity associated with an offload cue was similar, but not identical, to brain activity associated with a forget cue. We conclude by suggesting possible applications of the results to finding methods for improving intention offloading and avoiding memory failures.

Impact Statement

Cognitive offloading is one of the most crucial cognitive phenomena in daily life, with all of the technological advancements that allow for increased opportunities to store information on external devices rather than relying solely on internal memory. However, there remains a large number of unanswered questions about how different factors (e.g., aging, task type) influence strategic offloading decision and underlying neural correlates of cognitive offloading.

The thesis first used an offloading task, which evaluated not only whether factors (i.e. metacognitive advice, aging) facilitated or reduced offloading behaviour, but also whether the factors that influenced people resulted in offloading intention optimally. The results suggest that people do not always offload intentions optimally, and adapt intention offloading based on metacognitive judgment about unaided memory ability. This highlights the idea that metacognitive interventions can improve people's bias in strategic offloading decision. Moreover, the thesis shows that in some situations, older people had a reduced preference for offloading when their unaided memory ability was taken into account. A reduction in pro-reminder bias (i.e. bias towards offloading) in older adults could be explained by the resulting upward shift in confidence about unaided memory, relative to actual performance. This suggests metacognitive intervention can be an effective way to improve people's cognitive strategy usage across the lifespan.

Second, the thesis showed aversion to cognitive effort, along with metacognitive judgment, could play an important role in decisions related to offloading intention. The two cognitive processes, metacognitive judgment and

aversion to cognitive effort, were influenced by various intrinsic (i.e., aging) and extrinsic (i.e., PM task types, metacognitive advice) factors. This indicates that approaches to perform metacognitive intervention or improve offloading strategies should be tailored to varying individual and environmental conditions. Finally, the thesis suggests that the similarities of cognitive processes between cognitive offloading and directed forgetting sheds light on a potential research direction for investigating underlying neural correlates of cognitive offloading.

The findings presented in this thesis provide a fresh perspective on cognitive neuroscience in the area of understanding cognitive processes that are part of cognitive offloading such as metacognition, memory, and decision making. Crucially, the results help establish a theoretical account for the underlying mechanisms of cognitive offloading. Thus, the thesis delivers essential stepping stones to studying the psychological processes and neural correlates of cognitive offloading.

ACKNOWLEDGEMENTS

First, I would like to thank my supervisor, Prof. Sam Gilbert, for providing invaluable guidance and support throughout the research. I am extremely grateful for his help with various research difficulties as well as the thesis preparation. I would also like to thank my lab colleagues, Dr. Annika Boldt, Dr. Chiara Scarampi and Chhavi Sachdeva. The lovely time we shared in Room308 is the best part of this Ph.D. journey. Most importantly, it has been a great privilege to work and study in this lab and have their friendships. Besides, I would like to express my gratitude to my secondary supervisor, Leun Otten, for listening with empathy and for giving me a lot of research and career advice.

I am very much thankful to my family and friends for their love, understanding, and support. Without any of them, I would never accomplish this challenge. I must especially express my thanks to Lisa Liu, and Christine Chen for helping me through the toughest times. I am very grateful to all my past and present flatmates for a lot of joyful conversations and small trips we shared. I would also like to thank my land and country, Taiwan, for nurturing me, and providing me with a scholarship to complete the Ph.D. My Special thanks goes to Universe, for Her showers of blessings from the beginning to the end of this wonderful journey - I dedicate this work to Her.

UCL Research Paper Declaration Form: referencing the doctoral candidate's own published work(s)

Chapter 2

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- have been uploaded to a preprint server;
- are in submission to a peer-reviewed publication;
- have been published in a peer-reviewed publication, e.g. journal, textbook.

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2):	
a) Where was the work published? (e.g. journal name)	Journal of Experimental Psychology: General
b) Who published the work? (e.g. Elsevier/Oxford University Press):	American Psychological Association
c) When was the work published?	2020
d) Was the work subject to academic peer review?	Yes
e) Have you retained the copyright for the work?	No
[If no, please seek permission from the relevant publisher and check the box next to the below statement]:	
<input checked="" type="checkbox"/> <i>I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.</i>	
2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):	

a) Has the manuscript been uploaded to a preprint server? (e.g. medRxiv):	Please select.	If yes, which server? Click or tap here to enter text.	
b) Where is the work intended to be published? (e.g. names of journals that you are planning to submit to)	Click or tap here to enter text.		
c) List the manuscript's authors in the intended authorship order:	Click or tap here to enter text.		
d) Stage of publication	Please select.		
3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):			
Sam Gilbert was involved with the conceptualisation and data analysis of the entire paper. Pei-Chun Tsai was involved with the conceptualisation and data analysis of experiment 2.			
4. In which chapter(s) of your thesis can this material be found?			
Chapter 2			
5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work):			
Candidate:	Pei-Chun Tsai	Date:	17/11/2022
Supervisor/ Senior Author (where appropriate):	Sam Gilbert	Date:	17/11/2022

Chapter 3

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- have been uploaded to a preprint server;
- are in submission to a peer-reviewed publication;
- have been published in a peer-reviewed publication, e.g. journal, textbook.

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

6. For a research manuscript that has already been published (if not yet published, please skip to section 2):		
f) Where was the work published? (e.g. journal name)	Click or tap here to enter text.	
g) Who published the work? (e.g. Elsevier/Oxford University Press):	Click or tap here to enter text.	
h) When was the work published?	Click or tap to enter a date.	
i) Was the work subject to academic peer review?		
j) Have you retained the copyright for the work?	Please select.	
[If no, please seek permission from the relevant publisher and check the box next to the below statement]:		
<input checked="" type="checkbox"/> <i>I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.</i>		
7. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):		
e) Has the manuscript been uploaded to a preprint server? (e.g. medRxiv):	Yes	If yes, which server? PsyArXiv
f) Where is the work intended to be published? (e.g. names of journals that you are planning to submit to)	Click or tap here to enter text. Psychology and Aging	

g) List the manuscript's authors in the intended authorship order:	Click or tap here to enter text. Pei-Chun Tsai, Chiara Scarampi, Matthias Kliegel, and Sam J. Gilbert		
h) Stage of publication	Submitted		
8. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):			
Pei-Chun Tsai, and Sam Gilbert was involved with the conceptualisation and data analysis of the entire paper. Chiara Scarampi, and Matthias Kliegel were involved with conceptualisation.			
9. In which chapter(s) of your thesis can this material be found?			
Chapter 3			
10. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work):			
Candidate:	Pei-Chun Tsai	Date:	17/11/2022
Supervisor/ Senior Author (where appropriate):	Sam Gilbert	Date:	17/11/2022

Table of Contents

Abstract	3
Impact Statement.....	4
ACKNOWLEDGEMENTS.....	6
UCL Research Paper Declaration Form: referencing the doctoral candidate’s own published work(s)	7
Chapter 2	7
Chapter 3	9
1	16
General Introduction	16
1.1 Preamble	16
1.2 Prospective Memory	17
1.2.1 The Multiprocess Framework	21
1.2.2 Context	22
1.2.3 Prospective Memory across the Lifespan	24
1.3 Intention Offloading	25
1.3.1 Approach to Evaluate Intention Offloading	26
1.3.2 Metacognition	27
1.3.2.1 Metacognition in Prospective Memory	28
1.3.2.2 Contributing to Cognitive Offloading.....	31
1.3.3 Cognitive Load	33
1.3.4 Strategy Perseveration	35
1.3.5 Motivation	35
1.4 The Effects of Cognitive Offloading on Memory Abilities	36
1.4.1 Directed Forgetting.....	38
1.5 Neural Correlates of Cognitive Offloading	42
1.6 Issues to be Addressed in the Current Thesis	43
1.6.1 Research Questions	43

1.6.2	Overview of Experiments.....	44
2	46
	Experiment 1: Optimal Cognitive Offloading and Metacognitive Advice.....	46
2.1	Introduction.....	46
2.1.1	Aims.....	48
2.2	Method and Procedure.....	49
2.2.1	Participants.....	49
2.2.2	Method.....	50
2.2.3	Procedure.....	51
2.2.4	Dependent Measures.....	54
2.2.5	Optimal and actual indifference points:.....	55
2.3	Results.....	57
2.4	Discussion.....	62
2.4.1	Metacognitive Judgement and Reminder Bias.....	64
2.4.2	Avoiding Cognitive Effort.....	65
2.4.3	Limitations of this study.....	66
2.4.4	Practical implications and future research.....	67
3	70
	Experiment 2: Age Differences in Optimal Cognitive Offloading.....	70
3.1	Introduction.....	70
3.1.1	Prospective Memory and Aging.....	70
3.1.2	Metacognition and Aging.....	73
3.1.2.1	Metacognitive control and aging.....	75
3.1.3	Hypothesis.....	77
3.2	Method and Procedure.....	77
3.2.1	Participants.....	77
3.2.2	Method.....	78
3.2.3	Procedure.....	80
3.2.4	Dependent Measures.....	82

3.2.5	Exclusion Criteria	84
3.3	Results	85
3.3.1	Accuracy at the Delayed-intention Task	86
3.3.2	Reminder Setting Behaviour	86
3.3.3	Reminder Bias.....	88
3.3.4	Metacognitive Judgements.....	88
3.3.5	Additional, Non-preregistered Analyses	89
3.4	Discussion	91
3.4.1	Limitations of this study	95
3.5	Conclusion	97
4	98
	Experiment 3: Effects of Delay Length and Metacognition on Strategic Reminder setting in Time-based Prospective Memory	98
4.1	Introduction.....	98
4.2	Experiment 3a.....	102
4.2.1	Method	102
4.2.2	Procedure	104
4.2.3	Participants.....	105
	Exclusion Criteria:	106
4.2.4	Dependent Measures	107
4.2.5	Results	108
	4.2.5.1 Metacognitive Judgment and Reminder Usage	110
	4.2.5.2 Modification across Three Delays	114
	4.2.5.3 PM Performance Improvement	115
	4.2.5.4 Ongoing Task Performance	116
	4.2.5.5 Additional Non-preregistered Analyses	117
4.3	Experiment 3b	119
4.3.1	Method and Procedure.....	119
4.3.2	Participants.....	120
	Exclusion Criteria:	121

4.3.3	Dependent Measures	122
4.3.4	Results	122
4.3.4.1	PM Performance.....	123
4.3.4.2	Reminder Usage.	123
4.3.4.3	Metacognitive Judgement.	125
4.3.4.4	Metacognitive Judgment and Reminder Usage	126
4.3.4.5	Modification across Three Delays	128
4.3.4.6	PM Performance Improvement	129
4.3.4.7	Ongoing Task Performance	130
4.3.4.8	Additional, Non-preregistered Analyses	134
4.4	Discussion	135
4.4.1	Time Monitoring Behaviour.....	137
4.4.2	Reminder Benefit.....	139
4.4.3	Limitation and Future Research.....	140
4.4.4	Conclusion	142
5	143
	Experiment 4: The Neural Correlates of Cognitive offloading: Compared with Intentional Forgetting and Remembering.....	143
5.1	Introduction.....	143
5.2	Method and Procedure.....	147
5.2.1	Stimuli.....	147
5.2.2	Method	148
5.2.3	Participants.....	150
	Inclusion Criteria.....	151
5.2.4	Procedures.....	151
5.2.4.1	On-line phase	151
5.2.4.2	MRI Scan.....	151
5.2.4.3	Post-MRI Memory Test	152
5.2.5	fMRI Recording and Data Analysis	153
5.3	Results	154
5.3.1	Behavioural Results	154
5.3.1.1	Results for Position Test.....	154

5.3.1.2	Results for Recognition Test.....	156
5.3.2	fMRI Results.....	158
5.3.2.1	Forget vs. Remember Trials	158
5.3.2.2	Help vs. Remember trials.....	159
5.3.2.3	Forget vs. Help trials	159
5.4	Discussion	162
5.4.1	Saving-enhanced Memory Effect and Categorical Incompatibility	165
5.4.2	Limitation and Future Research.....	168
6	170
	General Discussion	170
6.1	Overview of the Findings	170
6.2	Cognitive Offloading in PM Tasks	171
6.3	Factors influencing cognitive offloading	172
6.3.1	Metacognition	172
6.3.2	Cognitive load	174
6.4	Neural Correlates of Cognitive Offloading	176
6.5	Beneficial Effects of Cognitive Offloading	177
6.6	Limitation and Suggestions for Future Research	178
6.7	Conclusions.....	181
	References.....	182

General Introduction

It is possible, of course, to jot them down on a notepad or something of the sort, but I prefer to trust my mind. It's a real pain to carry a pad around, and I have found that once I have jotted something down I tend to relax and forget it.

— Haruki Murakami, *So What Should I Write About*

1.1 Preamble

The study of cognition extends beyond activities within the brain. Instead, understanding how the mind works requires examining the contribution of external resources, including both body and environment. Our capability to adaptively incorporate internal cognitive processes with external resources may define us as successful cognitive agents in an intricate environment. For example, we do not need to know the entire layout of a city to walk from one location to another. Instead, we remember a few turning points and follow the streets connecting them (Clark, 1996). Another example is that basketball players adjust their speed of running to keep the same angle between their eyes and the ball in motion so that they do not need to intercept the ball with difficult calculation (Keijzer, 2002). Previous research (Clark & Chalmers, 1998; Heersmink & Sutton, 2020) proposed the term “active externalism” which refers to the active contribution of external cues to cognitive processes. Risko and Gilbert (2016) defined “cognitive offloading” as a skill or a strategy people use to incorporate physical action to create external tools simplifying the cognitive demands of a task. Cognitive offloading not only allows people to store information in-the-

world (Risko & Dunn, 2015) but also allows the offloaded information to modify people's behaviour later. A case in point is that people write down notes or set alarms which will remind them later of actions to complete. In everyday life, we often rely on external tools and resources to support our cognitive processes. Increased opportunities to save information in external stores were offered by the rapid advancement of technology taking place in our society such as smartphones, GPS devices, computers, and search engines (i.e. information could be searched instead of remembered; Brügger et al., 2019; Ferguson et al., 2015; Fisher et al., 2015; Herrmann et al., 1999; Marsh & Rajaram, 2019; Morrison & Richmond, 2020; Storm & Stone, 2015)

In all cases where people use an offloading strategy, it is common to assist in reducing the failure of prospective memory (PM) in everyday life (Henry et al., 2012; Jones et al., 2021). This introduction will provide a review of cognitive offloading literature in the memory domain, with an emphasis on PM. Before discussing cognitive offloading in PM (i.e. intention offloading), we address the PM processes and relevant contributing factors.

1.2 Prospective Memory

PM is defined as people remembering an intention that should be executed at a proper point in the future (Einstein & McDaniel, 1990; Brandimonte et al., 1996), such as buying a loaf of bread on the way home. PM has been proposed to involve several components that starts with planning an intended action, followed by monitoring the intention consciously or unconsciously until the action is carried out at the proper time (McDaniel & Einstein, 2011). For example, people have the idea to

buy a loaf of bread on the way home when they go to work in the morning. They plan to choose the route passing a bakery when they come home. This intention may be stored in their memory when they are engaging in other activities, such as working in the office. The cues or targets can be monitored autonomously or consciously during this period of time. For instance, they may have rehearsals of this intention or some words like bakery may remind them of this future action when they come across it reading magazines. At last, when they leave the office, they retrieve this intention, choose the correct route and buy the bread. To perform the PM task, people need to remember the prospective PM component (i.e. remembering something should be done) and to remember the retrospective PM component (i.e. remembering what should be done). The failure of PM, however, seems not uncommon in our everyday life. People sometimes remember the delayed intention, buying bread, earlier but forget it when they leave the office. They finally think of it at home although it is too late. Therefore, self-initiated retrieval plays an important role in achieving PM tasks since people need to remind themselves of an intention at the proper time when it shall be fulfilled.

PM tasks can be categorised by distinguishing time-based and event-based PM tasks (Einstein & McDaniel, 1990; Harris, 1984; Brandimonte et al., 1996). In event-based tasks, people fulfill the delayed intention when a particular event happens such as the example of buying bread on the way home while a time-based task requires the intended action to be carried out at a designated time or after a specific interval such as buying bread at 4 p.m. The difference between event-based and time-based PM has not been clearly defined. Some researchers argued that time could be simply thought of as a less forceful cue than an event cue (Uttl, 2008). Moreover, the time cue could be replaced by a secondary event cue such as buying

bread after the meeting which will finish at 4 p.m. Other research suggested event- and time-based PM might involve diverse monitoring processes such as time estimation and time monitoring in time-based PM (Conte & McBride, 2018; Harris & Wilkins, 1982; Kvavilashvili & Fisher, 2007; Waldum & McDaniel, 2016).

Studies on PM may be carried out in laboratory or naturalistic settings. In laboratory settings, the paradigm which has been used broadly is the one proposed by Einstein and McDaniel (1990). In this paradigm, participants were usually asked to perform an ongoing task while they waited for the appropriate time to complete a delayed intention, which simulates daily PM tasks which people perform when they are doing other activities. For example, participants were presented with a set of words for immediate recall on each trial and also told to make a special response if they saw a specific word. The measure of PM was whether they remembered to make this special response, and the PM task was embedded within the short-term memory task. This is like remembering to buy bread when passing a bakery instead of making a special trip for the bread. An ongoing task like the short-term memory task represents the daily activities people need to carry out and impedes rehearsal of the delayed intention. Moreover, some laboratory or clinical approaches have been developed to mimic daily PM situations such as the virtual week task (Rendell & Craik, 2000) and the hotel test (Manly et al., 2002). Results found in the basic laboratory tasks are consistent with the more realistic tasks. For example, the studies of both experimental settings found that the age difference was reduced by decreased demand on self-initiated retrieval (Kliegel, Jäger, et al., 2008; McDaniel et al., 2009; Rose et al., 2010).

In contrast to laboratory studies, the naturalistic setting examines participants'

behavior in daily life through task-performance and observational assessment. Experimenters may ask participants to perform instructed activities such as phoning them at a specific time in the future (Moscovitch, 1982), or self-assigned intentions (e.g., Haas et al., 2022). Mobile technology has been used to collect data for these observational studies by regularly recording participants' daily conversation to find how frequently they mention failing PM tasks (Haas 2022).

Some researchers indicated that retrieving intentions from long-term memory was an essential part of the PM (Brandimonte M. A. et al., 2001; Graf & Uttl, 2001). PM tasks should occur infrequently to engage participants in ongoing tasks, similar to real-life PM tasks where people focus on their everyday activities. Moreover, the intention may fade from attention and need to be brought back to awareness at the appropriate time in the infrequent PM tasks, whereas participants think about the intention continually in vigilance tasks. Therefore, PM tasks in the laboratory setting are, therefore, considered vigilance tasks (i.e. involving explicit rehearsal of a delayed intention over a short duration) seeing as the PM trials are usually frequent (Brandimonte et al., 2001; Graf & Uttl, 2001). Some studies (Brandimonte et al., 2001; McDaniel et al., 1998) proposed that PM is based more on autonomic retrieval while vigilance relies more on monitoring of top-down control. However, Gilbert(2011) found rostral prefrontal cortex was associated with remembering intentions in PM tasks lasting few seconds, which was consistent with the results of previous studies evaluating daily PM tasks (Burgess, 2000; Uretzky & Gilboa, 2010). It could be rational to speculate about some overlapping cognitive processes across different timescales of PM task (Gilbert 2015a).

1.2.1 The Multiprocess Framework

Given self-initiated retrieval is highly required in PM, the monitoring processes have been characterised in existing theoretical frameworks. An early approach, the test-wait-test-exit model (Harris & Wilkins, 1982), on time-based PM defined the monitoring processes descriptively. It suggested people would repeatedly check whether it is the appropriate time to fulfill the intention. If it was not, they would wait until checking (i.e. time-monitoring) again. In the end, people performed the delayed intention when the designated time was reached. More recently, a theoretical approach, the multiprocess framework view (MPV; Einstein et al., 2005; McDaniel et al., 2004; McDaniel & Einstein, 2000), developed in the context of event-based PM, proposes that monitoring processes are mediated by dual pathways (for a review see McDaniel et al., 2015). The top-down pathway accounts for attention control conditions such as actively rehearsing the delayed intention while the bottom-up pathway supports spontaneous retrieval. For example, to complete a PM task, like buying bread at 4 p.m., people encode the target time to be related to the intention, which attaches intention-based significance to the target time, and influences how the target time will be processed later. People then may spontaneously notice the time when their eyes scan a clock on a computer screen, and the top-down control is engaged to become aware of the PM intention as well as manage the required action responses. On the other hand, people may actively rehearse the intention and check the clock regularly to avoid forgetting. The extent to which people are biased towards either top-down or bottom-up pathway would depend on the features of PM and ongoing tasks. For example, the spontaneous retrieval can be encouraged by overlapping processes between PM and ongoing tasks (focality; e.g., both tasks require encoding semantic information; see next

section for more details; McDaniel & Einstein, 2000).

Recently, MPV has been extended into the dynamic multiprocess view (DMPV), which emphasises that the dual pathways operate simultaneously during PM processes and affect the involvement of each other mutually (Scullin et al., 2013; Shelton & Scullin, 2017). Continuing with the previous example of “buying bread,” if people spontaneously retrieve the intention at 3 p.m., they may then regularly monitor the time until going to the bakery. Conversely, actively rehearsing the intention facilitates a spontaneous retrieval later. Another amended argument of DMPV is that bottom-up and top-down processes do not only contribute to the later retrieval phase as described in the MPV, but also to the earlier stages such as intention formation and intention retention. For example, although intention formation is often considered a strategic process, some studies have revealed that people from time to time form intentions automatically without carefully thinking about how and where they would perform the intention (Holbrook & Dismukes, 2009; Scullin et al., 2018).

1.2.2 Context

Context is an important contributing factor for PM (R.E. Smith et al., 2017) on both an operational and a theoretical level. On an operational level, PM has been studied within various contexts such as laboratory versus naturalistic settings, time-based versus event-based, and internal memory versus memory aids. The term "context" can refer to various processes. First, context can be defined as the different conditions or settings where a PM task is carried out as well as a signal indicating when a PM task can be performed. Moreover, Environmental aids are a

type of context which can influence whether people could successfully complete a PM task. External cues usually serve as memory aids to improve the performance of PM tasks such as a noticeable clock for a time-based task. Also, focality of a PM target which takes both the PM target and context into account, impacts intention retrieval and performance (Einstein et al. 2005). Focality has been defined as the extent to which processing the PM target overlaps with the processing of the ongoing task where the PM task is embedded. For example, when participants perform an ongoing, lexical task where they decide whether a word is a real word, and are asked to respond to a word for a PM task, the PM task is considered “focal” within the ongoing task.

On a theoretical level, Preparatory Attentional and Memory Process theory (R. E. Smith, 2003) suggested that establishing an internal context is essential for PM processes. For instance, buying bread at 4 p.m. means the action should be delayed until the proper context, 4p.m., occurs. The preparatory attention processes (i.e. internal context) are engaged in matching the intention and the retrieval, but the constant engagement is not necessary. R. E. Smith et al. (2017) indicated that the resources required by preparatory attention processes are directly related to the proximity to the PM targets in a predictable environment. For the prior example, we prepare stopping at the bakery when the time is close to 4 p.m. rather than the whole day. R. E. Smith (2003) found that the performance on the ongoing tasks where no PM task was embedded was better (i.e. improved accuracy or shorter reaction time) than the ongoing tasks with PM tasks. The cost effect of PM tasks on ongoing activities could indicate an allocation of attention resources to PM rather than ongoing activities (e.g. Smith et al., 2017; Smith & Loft, 2014).

1.2.3 Prospective Memory across the Lifespan

Previous evidence from laboratory studies has found the earliest age when children show the ability to remember delayed intention might be at the age of 3 years (Causey & Bjorklund, 2014; Kelly et al., 2018). PM develops considerably at age 7 and 10 (Kliegel et al., 2013; R.E. Smith et al., 2010). In terms of the prospective PM component, research suggests that executive functions, including working memory, inhibition, shifting, and monitoring, plays a crucial role in the development of PM ability during childhood. On the other hand, evidence from the studies evaluating the development of memory aspects (i.e. the retrospective PM component) has provided contradicting results. Kliegel, Mackinlay, et al.(2008) found there was no age differences in the memory aspects of PM between children and adults while R. E. Smith et al.(2010) suggested adults had better performance of retrospective PM component than children.

In terms of ageing effects, research has shown that young adults often outperform older adults on PM tasks in laboratory settings (Henry et al., 2004; Kliegel et al., 2016). However, older adults often perform better than younger adults considerably in naturalistic approaches (Henry et al., 2004; Kliegel et al., 2016) where studies were carried out by asking participants to call the experimenters at specific times or send postcards. The pattern including the age-related loss in laboratory tasks and age-related benefits in naturalistic tasks has been described as a "paradoxical" effect (Rendell & Craik, 2000). Although some research suggested that this age PM paradox could be explained by the broader use of reminders in naturalistic settings (Dobbs & Reeves, 1996), the argument was not supported by more recent investigation (Aberle et al., 2010; Ihle et al., 2012; Phillips et al., 2008).

In line with those studies, an investigation on everyday PM experiences in real life (Hertzog et al., 2019) found older people usually expected themselves to remember important PM tasks, relying highly on unaided memory, and did not view external strategies as a compensatory aid for memory decline.

1.3 Intention Offloading

The rapid advancement of technology offers increased opportunities to offload intentions on external devices such as set reminders on mobile calendar. Given that the failures of PM contribute to the majority of daily memory problems (Kliegel & Martin, 2003), but people do not always offload delayed intentions to support PM, a relevant question is raised: how is intention offloading influenced by a variety of dimensions including the nature of individual PM and ongoing tasks, as well as the characteristics of these individuals? Since both internal processes and environmental structures contribute to intention offloading, numerous associated intrinsic (e.g., ageing) and extrinsic (e.g. the characteristics of PM and ongoing tasks) factors could be manipulated or held as keys to understanding the various cognitive capacities underlying the processes of intention offloading. More importantly, evaluating the adjustment of intention offloading allows us to examine the interaction and integration of cognitive capacities (e.g. PM, metacognition, decision making), which have seldom been investigated as a whole. In addition to contributing to the above academic progress, the research on intention offloading could provide possible methods to reduce PM failures, and help people to offload delayed intention more efficiently.

1.3.1 Approach to Evaluate Intention Offloading

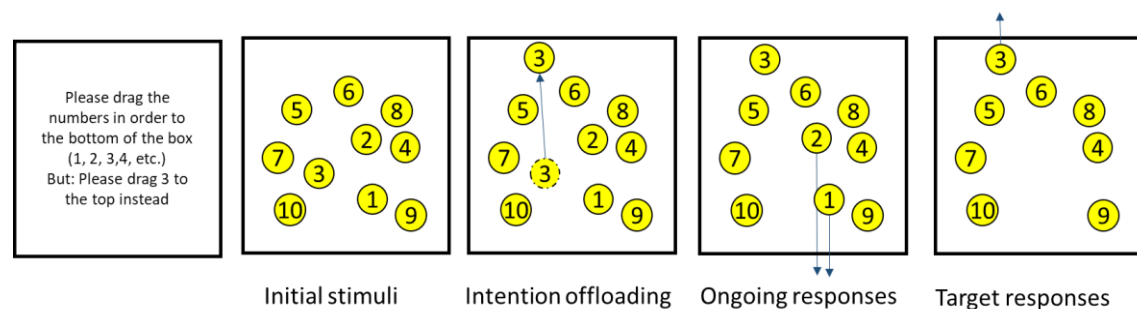
Previous studies on intention offloading have evaluated whether cognitive offloading improves PM performance by comparing the conditions where reminders are provided or not (Guajardo & Best, 2000; Henry et al., 2012; Kliegel & Jäger, 2007; Lourenço & Maylor, 2015; Mahy et al., 2018; Ryder et al., 2022). Although the consequences of offloading have been examined, a question of how offloading intentions are triggered remains. In order to investigate this question, experiments allowing participants to set reminders freely could examine the extent to which different factors affect intention offloading.

Gilbert (2015a, 2015b) proposed an intention offloading task to investigate how people decide whether to use external reminders or to depend on their own memory in PM tasks. See Figure 1.1 for a schematic illustration of the task. Participants were instructed to drag ten circles successively to the bottom of the square in a numerical sequence, in order to make the circles disappear. In each trial, most circles were dragged to the bottom edge except some target circles. At the beginning of each trial, the participants were instructed as to which target circles should be dragged to other indicated edges when they reached them in the sequence (e.g. "please drag 3 to the top instead"). Then, an offloading strategy was then explained to the participants, which consisted of setting reminders by dragging the target circle next to the instructed edge at the beginning of each trial. The location of the target circle represented the delayed intention and provided a perceptual cue to indicate which border the circle should be dragged to. This can be seen as similar to leaving an item by the front door to remind you to bring it when you leave home. Participants could choose freely between depending on unaided memory or the offloading strategy. In

spite of a relatively short retention time of delayed intention in this task, in a web-based study (Gilbert, 2015a), the memory performance of this task can predict the achieve rate of a real-world delayed intention task of visiting an indicated web link on three specific days in one week.

Figure 1.1

Schematic Illustration of the Intention Offloading Task



Note. A participant was told that Circle 3 is the target circle that should be dragged to the top. The delayed intention could be offloaded by dragging Circle 3 next to the top. After that, Circles 1 and 2 were dragged to the bottom before Circle 3 would be reached (ongoing responses), and the circles would disappear beyond the border of the square. Then Circle 3 was dragged to the top edge to fulfil the delayed intention task.

1.3.2 Metacognition

1.3.2.1 Each time we form an intention, we need to decide whether to remember it or set an external reminder. How do individuals make these decisions? This could be explained by an influence of metacognition, our ability to monitor and control our own cognitive processes and abilities (Dunlosky & Metcalfe, 2008; Flavell, 1979; Fleming, Dolan, et al., 2012; Koriat, 2007; Nelson & Narens, 1990). Metacognition, one of the cognitive processes

involved in intention offloading, has been defined broadly as "thinking about thinking". For example, people with efficient metacognition report higher confidence when giving correct answers than otherwise (J. D. Smith et al., 2003). Metacognition helps people to understand how their cognitive abilities work, learn different strategies and assess the competency of these strategies. Moreover, Nelson and Narens (1990) provided an early classification of metacognition as consisting of metacognitive monitoring and metacognitive control.

Metacognitive monitoring involves assessing the current state of a cognitive activity. This informs metacognitive control to drive people to adaptively change behaviour (e.g., Dunlosky & Metcalfe, 2008; Nelson & Narens, 1990). On the other hand, cognitive control refers to the intentional selection of thoughts, emotions, and behaviour based on current task demands and social context, and the concomitant suppression of inappropriate habitual actions. Although metacognitive control is a subcategory of cognitive control, cognitive control can operate independently of metacognition in some situations. For example, cognitive control can be automatically engaged in response to salient or threatening stimuli, without the need for metacognitive monitoring or judgment.

Metacognition in Prospective Memory

Kuhlmann (2019) described different metacognitive monitoring and control processes based on three PM phases: intention formation, intention retention and

intention retrieval. Continuing with the previous example of buying bread, when intention formation occurs, you predict the likelihood of remembering to buy bread on the way home, an example of metacognitive monitoring. You then plan to walk home on the road past the bakery, an instance of metacognitive control. During the day, you might self-initiate rehearsing the intention, again metacognitive control, and assess whether you will remember to do it later, an instance of metacognitive monitoring. If you fail to buy bread at the end, next time you may believe you cannot remember, which is a form of metacognitive monitoring, and decide to set a reminder on the smartphone, falling under metacognitive control. As can be seen, the outcome of metacognitive monitoring triggers metacognitive control processes, which modifies PM processes. It further reciprocally leads to an update of metacognitive monitoring, which indicates an intrinsic interrelationship of metacognitive monitoring and control.

Theoretical accounts of PM have suggested metacognition plays an important role (Kuhlmann 2019). The Preparatory Attentional and Memory Process theory suggested that assessing the association between contexts and PM tasks is a form of metacognitive monitoring, which guides metacognitive control to allocate attention to ongoing tasks versus PM tasks (R. E. Smith & Skinner, 2019). The multiprocess view and the dynamic multiprocess view assumes that metacognitive awareness of the cognitive demand of PM tasks guides metacognitive control processes, involving the extent to which people rely on the two pathways of automatic intention retrieval versus strategic processes (Einstein et al., 2005; Shelton & Scullin, 2017).

Various approaches have been utilised to assess metacognitive monitoring on PM performance, such as predicting PM performance (Cauvin et al., 2019; Meeks et

al., 2007; Schnitzspahn, Zeintl, et al., 2011), or through questionnaires (Rummel et al., 2019). Although Devolder et al. (1990) found that people predicted remembering to make phone calls more often than they actually remembered to make them, recent research shows that people have competent but imperfect metacognitive judgements about their PM ability and tend to underestimate their ability (see Kuhlmann, 2019 for a review). By further delineating situations, Schnitzspahn, Zeintl, et al. (2011) showed people were more confident about recalling the intention (i.e. the retrospective PM component) than remembering to do it at the appropriate time (i.e. the prospective PM component), which shows metacognition can differentially monitor both components.

On the other hand, findings on whether metacognition is sensitive to the cognitive demand of PM task could be varied depending on the factors which are manipulated. For example, research showed that metacognitive prediction is sensitive to memory load (i.e. one versus three PM targets; Cherkaoui & Gilbert, 2017; Gilbert, 2015b; Scarampi & Gilbert, 2021), but not to focality (Hicks et al., 2017). Further studies are needed to understand underlying mechanisms of how metacognition becomes aware of different modifications to task demand.

In terms of the metacognitive control in PM tasks, an investigation (Rummel & Meiser, 2013) showed the extent to which participants' strategies of attention allocation could be manipulated by explicit suggestions via modification of the participants' metacognitive judgment. Participants were given varied fake information designed to lead them to form different metacognitive beliefs about attentional demands. In the first experiment, people in the metacognitive-information condition were told that a red-coloured word before PM instructions

could make the task easier. In the second experiment, people were told that the pseudo-primers preceding PM targets make the task easier or harder under easy and difficult PM expectations conditions respectively. The results revealed that they allocated their attention more in line with these metacognitive judgment than the factual demands.

1.3.2.2 Contributing to Cognitive Offloading

Given how metacognition could monitor cognitive ability and guide people to adjust their strategies, it is not surprising that previous research supports the notion of metacognition playing an important role in cognitive offloading. Risko and Dunn (2015) used a short-term memory task to reveal that participants sometimes chose the offloading strategy although their performance was already maximised without utilization of the strategy. A possible explanation was that metacognitive bias would make people choose an external strategy because they believe they would perform poorly without offloading strategy even though it offered no advantage. In a Gilbert's study (2015a), half of the participants were interrupted by arithmetic questions when performing the intention offloading task while the other participants only performed the task. Participants were instructed to remember either one or three targets. Results showed participants set reminders more often when they were interrupted by arithmetic questions in the task as well as when they had three targets to remember instead of one. The interruption and increased number of targets caused higher memory load and compromised the performance of delayed intention task. If people are knowledgeable about the memory tasks and strategies (i.e. the task of more targets would be more challenging), it is possible that their confidence is lower due to perceiving the task as more difficult. However, it is also

possible to suppose that people have basic metacognitive insight about their performance and the insight about the reduced performance motivates offloading delayed intention (Arango-Muñoz, 2013; Weis & Wiese, 2019).

More evidence converging on offloading behaviour being driven by metacognitive judgments has demonstrated that metacognitive judgement about performance contributes to the extent to which people offload information, regardless of how well they actually perform (Gilbert 2015b, Dunn & Risko 2016). A study of Gilbert (2015b, Experiment 1) required participants to predict what percentage of target responses when performing the intention offloading task. In Phase 1, participants performed the task using internal memory while they were allowed to use reminders in Phase 2. In one experiment (Gilbert, 2015b, Experiment 1a), the reminder usage in Phase 2 was predicted by the Phase 1 predictions, which indicates people tended to use more reminders if they believed they performed poorly using internal memory. However, the first-phase predictions was not related with the Phase 1 performance, which demonstrates the adjustment of intention offloading at least partially is attributed to metacognitive judgment. Though participants were instructed how to set reminders in the study, people have to initiate reminders themselves in daily life. It remains a question whether metacognitive processes has the same impact on intention offloading when people spontaneously initiate an offloading strategy. In a study of Boldt and Gilbert (2019), whether participants were provided with instructions on setting reminders was manipulated between two groups. Some of the participants who were not given the reminder instruction spontaneously came up with the strategy. Although the spontaneous group produced less offloaded intentions, results were found in both groups, which showed metacognitive judgment contributed to intention offloading,

regardless of memory performance.

In addition to the effects of metacognitive judgment about memory performance, the interaction between metacognitive processes and other cognitive processes such as cognitive effort and motivation (see the sections below) also impact intention offloading. For example, people might write down a checklist on a smartphone memo because they think it is important. The awareness of the importance, namely the interplay between metacognition and motivation processes, triggers and modulates intention offloading. Having discussed the pivotal role of metacognitive processes in modifying intention offloading, it is reasonable to suggest that improving metacognitive monitoring can lead to offloading intentions more efficiently and optimally. To further study the possible benefits of such intervention programmes, Experiment 1 investigated whether metacognitive advice would optimise the use of offloading strategies.

1.3.3 Cognitive Load

Cognitive offloading could save people from using internal memory to remember intentions to a certain extent. It is reasonable to consider the magnitude of cognitive effort required by a PM task as an important factor affecting cognitive offloading. The term, cognitive effort, is difficult to operationalise, which is analogised to a more concrete example of physical effort such as carrying a heavy object (Shenhav et al., 2017). People can use a tool like a trailer to avoid physical effort. Would this be the case with cognitive effort? The minimum memory hypothesis (Wilson, 2002) suggests that people tend to choose external resources to reduce the use of internal representations. While this hypothesis highlights the

preference for external strategy insofar as they can be applied, other theoretical accounts (Dreisbach & Fischer, 2015; Kool et al., 2010; Kurzban, 2016) propose the conservation of cognitive effort as the main concern regardless of internal memory or external resources. Previous empirical evidence for the avoidance of cognitive effort is mixed. Some research suggested individuals might choose a less optimal strategy which minimises cognitive effort (Kool et al., 2010), and cognitive effort could trigger aversive signals. However, some studies have proposed arguments against the postulation of cognitive effort aversion. Evidence, showing people prefer more cognitively challenging tasks, and value the outcome more when more effort is exerted to gain it, has suggested cognitive effort can be perceived as an internal reward rather than a cost (for a review see Inzlicht et al., 2018; Norton et al., 2012; Olivola, 2018). The mechanism by which cognitive effort differentially becomes a motivating or aversive factor remains unclear, and needs more future investigation (Inzlicht et al., 2018).

In the context of having inner bias to reduce cognitive effort, one possible explanation for the bias is that internal cognitive resources are finite (Baumeister et al., 2007). The theoretical account of limited resources, however, has been challenged by some previous studies, arguing that the supposedly limited cognitive resources has not been clearly defined, and the tasks used to evaluate the account have not been independently validated (Hagger et al., 2010; Lurquin & Miyake, 2017). Critically, another constraint derived from the limited resources is the lack of the ability to performing multiple tasks simultaneously when the same cognitive processes are involved in different tasks. Dedicating the resources in one task means losing an opportunity to use it on another task during the same period of time. Thus, an alternative account could attribute the avoidance of cognitive effort to avert the

cost of lost opportunities (Kurzban et al., 2013).

1.3.4 Strategy Perseveration

Previous habits may influence people's choices. Research has shown that there is a cost to changing strategies; people perform more poorly when they use different strategies than the ones they used before (Lemaire & Lecacheur, 2010; Taillan et al., 2015). Gilbert (2015a) using intention offloading task revealed participants tended to either always set reminders or utilising unaided memory. In Scarampi and Gilbert's study (2020, Experiment 2), participants performing an intention offloading task were first forced to use reminders or use their own memory. In a later phase, they could choose whether to use reminders. It was found that participants tended to choose the strategy they had used in the earlier phase. Therefore, the decision of whether to use an intention offloading strategy is influenced by prior experiences or cognitive intervention of offloading strategies.

1.3.5 Motivation

Memory literature has shown that people are usually willing to assign more time to remember high-value information, which lead them to have better recall for it (Castel, Murayama, et al., 2013; Castel, Rhodes, et al., 2013). This premise was defined by Murphy and Castel (2020) as the term responsible remembering. In their later study (2021), participants were assigned to remember some items of a list with hypothetical friends remembering the remaining items. They found that the recall proportion of each item was correlated with the instructions; in other words, participants recalled better the items which were assigned to them. Additionally, if

asked to imagine going camping, and to assess how important each item is when they studied the words, participants had better recall for what they deemed important items. This indicates motivation may play a pivotal role in the memory strategy involving prioritising more important targets. In terms of whether importance or consequences would influence intention offloading, a study of Dupont et al. (2022) using the intention offloading task demonstrated that participants tended to offload intentions associated with higher monetary reward compared to lower monetary reward.

On the other hand, each individual assesses the importance of a delayed intention and offloading strategies subjectively and differentially. For example, risk averters may choose offloading strategy as it can prevent the information from being lost and thus is a safer choice to avoid risks. In addition to the individual-specific estimates, environmental and social factors may contribute to the value processing of intentions as well as intention offloading. For example, Risko and Dunn (2015) showed the social presence of an experimenter enhanced participants' reliance on the offloading strategy, which suggests participants offloaded more intentions when being motivated by the presence of an experimenter to obtain better performance.

1.4 The Effects of Cognitive Offloading on Memory Abilities

Having discussed the possible cognitive processes involved in intention offloading, a relevant question should be considered next: what are the subsequent effects of cognitive offloading? These effects can be divided between cost and benefit, depending on the information the effect acts on. Storm and Stone (2015)

proposed offloading certain information can improve the memory for other information. In the study, having studied a word list, participants were instructed either to save or not save the list. Then, they studied the second list, and recalled the two lists separately. Participants saving the first list recalled the second list better than the participants that had to remember both lists. Runge et al. (2019) replicated the experiment and found that the benefits of offloading memory onto outside sources are not limited to memory performance. Instead, benefits also include improvement in the performance of subsequent, unrelated tasks.

On the other hand, offloading information could reduce the individual's memory for the offloaded information. Using the same two-list experimental paradigm, Runge et al. (2020) showed that participants who saved a word list, but did not know in advance that they would not be actually allowed to restudy it before testing had worse recall compared with not saving the list. Research supporting the cost effect of cognitive offloading has evaluated various memory processes, some of which are more naturalistic, such as taking photographs in a museum tour or using GPS when driving (Eskritt & Ma, 2013; Grinschgl et al., 2021; Henkel, 2014; Kang, 2022; Lurie & Westerman, 2021; Sparrow et al., 2011; Tamir et al., 2018). Critically, given that plenty of research has demonstrated the cost effect of cognitive offloading, two relevant questions are raised which is worth careful evaluation. First, what will be the consequence if the external aids fail? A study of Dupont et al. (2022) demonstrated participants tended to offload intentions associated with higher values. This implies that having offloaded important information, people may only remember the low-value information if the external aids are removed. The second question is whether the memory reduction caused by using an offloading strategy will lead to after-effects on memory abilities. Scarampi and Gilbert (2020, Experiment 1) showed using

an offloading strategy in the first phase did not significantly influence subsequent unaided memory ability. However, the effects of long-term usage of offloading strategies require further investigation.

1.4.1 Directed Forgetting

Storm and Stone (2015) postulated that the cost and benefit of cognitive offloading is an adaptive reallocation of cognitive resources. The theoretic account and experimental paradigm were inspired by the directed forgetting literature. Although forgetting has been often thought of as a memory failure, Bjork (1989) indicated that the effects of the forgetting processes involve pruning unrelated or outdated information and facilitating the learning of relevant items, which makes memory processes more efficient by avoiding interference. Forgetting is proposed to be more of an adaptive strategy as opposed to an error in memory. Previous directed forgetting research has principally used two paradigms (see Figure 1.2), item-method directed forgetting (IMDF) and list-method directed forgetting (LMDF). In IMDF studies, participants were first required to remember an item (e.g. a word, a picture), and then instructed to forget or remember by a cue following the item. After being shown a number of items, participants took a final memory test. IMDF studies have shown the memory performance for to-be-forgotten items were worse than to-be-remembered items. In LMDF research, participants were presented with a list of items, and the cue was given after the list. A typical trial of LMDF paradigm (for a review see Sahakyan et al., 2013) might consist of two or three lists with participants taking a memory test at the end of the trial. Results show the cost and benefits of directed forgetting: participants instructed to forget a previously studied list and remember a new one displayed diminished memory for the previous list and

memory enhancement of the second list. In general, when the memory test was a recognition test, IMDF studies found directed forgetting has an impact on the memory performance, whereas LMDF studies did not (Anderson & Hulbert, 2021; Bjork, 1989; Spitzer, 2014). However, when participants were required to recall freely in the memory test, directed forgetting affected the performance in both IMDF and LMDF studies (Anderson & Hanslmayr, 2014; Anderson & Hulbert, 2021; Bjork, 1989; Spitzer, 2014).

Previous research has proposed a number of theoretical accounts for mechanisms underlying directed forgetting such as differential rehearsal (Basden et al., 1993; Bjork, 1989), retrieval inhibition (Geiselman et al., 1983), contextual change (Sahakyan & Kelley, 2002), and inhibition of episodic memory (Racsomány & Conway, 2006). The diverse mechanisms may exist simultaneously, but manifest to varying extent depending on the specific task, such as item method versus list method, and encoding into long-term memory (for a review see Anderson & Hanslmayr, 2014; Anderson & Hulbert, 2021). Research has suggested that if the information has been encoded into long-term memory, directed forgetting is more likely to involve retrieval inhibition (Bjork, 1989). Formed representations implanted in long-term memory are likely inhibited rather than completely erased. On the other hand, if the information has not been encoded, passive decay (i.e. drop out of the rehearsal list) or active suppression of rehearsal might contribute to the effect of directed forgetting (Festini & Reuter-Lorenz, 2017; Zacks et al., 1996).

IMDF literature has focused more on differential rehearsal explaining the effects of directed forgetting: the to-be-forgotten items are exempt from further rehearsal while the to-be-remembered items continue to be rehearsed (Anderson &

Hanslmayr, 2014b; Tan et al., 2020). Recent LMDF studies highlight the role of episodic memory inhibition to a greater extent, as many experiments have revealed that the forget cues may induce inhibitory processes, which impair the performance of a subsequent task requiring the engagement of inhibitory processes (Fawcett & Taylor, 2008, 2010).

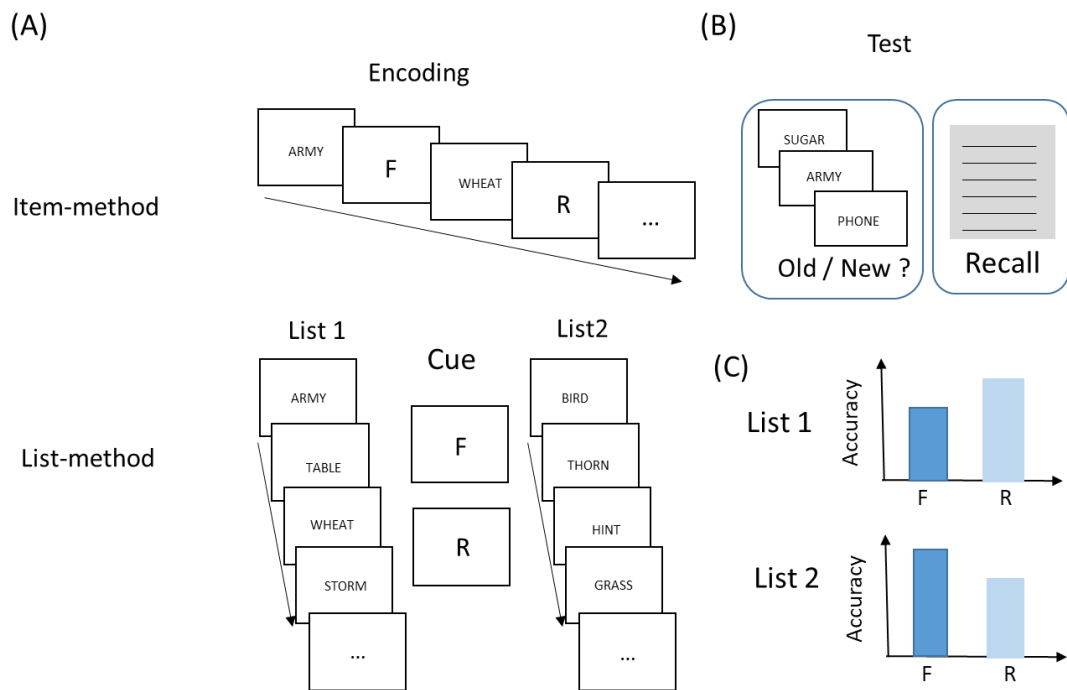
On the other hand, retrieval inhibition and context change has been more broadly discussed in LMDF studies (Anderson & Hanslmayr, 2014, Kliegl et al. 2020). The retrieval-inhibition account proposes that an active inhibitory processes elicited by cueing people to forget the pre-cue information reduce the accessibility of the information (Geiselman et al., 1983). With regard to the account of contextual change, given that the target list in LMDF could be encoded as representations of temporal context, the context of the list could be changed to prompt directed forgetting (Sahakyan & Delaney, 2005; Sahakyan & Kelley, 2002). Sahakyan and Kelley (2002) required participants to imagine an environment diversely when different lists were encoded, proposing that the mental context elicit a similar effect as directed forgetting.

LMDF research has revealed experimental evidence supporting the argument that the pre-cue suppression and the post-cue enhancement, namely the cost and benefit effects of directed forgetting, should not be explained by a single mechanism(for a review see Anderson & Hanslmayr, 2014). There were two possible theoretical accounts for the post-cue enhancement: a) cognitive resources freed by offloading information could enhance any encoding to come b) offloading information might benefit subsequent memory ability by reducing interference at recall. Sahakyan (2004) found evidence of better recall for a post-cue list, but Kliegel

et al. (2013) and Kliegel et al. (2020) failed to find post-cue enhancement. To evaluate the conflicting results, Pastötter and colleagues (2012) emphasised another factor, list output order, finding that reliable post-cue enhancement arose only when post-cue information was recalled first. It suggests re-exposure to pre-cue information (testing it before the post-cue information) might reinstate proactive interference, leading to a reduction of the enhancement effect. These results might be more in line with the second account, seeing as the freed cognitive resources account would predict that the output order should not influence the extent to which the freed resources boost the recall of post-cue information. However, more research is needed to clarify the underlying mechanisms of post-cue enhancement.

Figure 1.2

Schematic Illustration of the Item and List-methods for studying directed forgetting



Note. (A) In IMDF, each item was followed by a cue, and participants were tested all items at the end whereas in LMDF, participants first studied an entire list, followed by a cue, and then studied a second list. After the second list, the memory for the two lists were tested. (B) The memory tests in IMDF and LMDF might be a recognition test or a recall test. (C) Typical behaviour results found in list-method showed the effects of F-cue including memory reduction of List 1 and memory enhancement of List 2.

1.5 Neural Correlates of Cognitive Offloading

The similar cost and benefits of cognitive offloading and directed forgetting lead to a crucial question of whether cognitive offloading acts as a variant of intentional forgetting. Very few studies have investigated the extent to which subsequent brain activities incurred by cognitive offloading resemble intentional forgetting. Runge et al. (2021) evaluated the electrophysiological (EEG) oscillations of cognitive offloading

and did not find reduced alpha power and alpha phase synchrony, which has been related to the effects of intentional forgetting, during the encoding of the second word list. Further exploration is necessary in order to fully understand the mechanisms of cognitive offloading, and contributing factors such as encoding or retrieval.

1.6 Issues to be Addressed in the Current Thesis

1.6.1 Research Questions

The sections above review the literature on cognitive offloading, especially intention offloading. The psychological processes and neural mechanism of cognitive offloading are unclear due to the lack of direct empirical evidence revealing how individuals with intrinsic factors, such as ageing or brain injury, manage to exploit these compensatory strategies, and how extrinsic factors, such as metacognitive intervention or the characteristics of PM tasks, influence various cognitive processes (e.g. metacognition, cognitive effort) to adaptively alter intention offloading. There is also very little neuroimaging evidence regarding the brain activities related to cognitive offloading.

The current thesis aims to address the following two research questions about intention offloading: a) how is intention offloading influenced by different factors? b) what are the neural correlates of offloading information? To answer the first question, three behavioural experiments were designed in which the impact of intrinsic and extrinsic factors, including metacognitive intervention (see Experiment 1), ageing (see Experiment 2), and task type (see Experiment 3), was investigated

respectively. To answer the second question, fMRI was employed to compare the brain activities between the conditions where participants were cued to forget, remember or offload information (see Experiment 4).

1.6.2 Overview of Experiments

Experiment 1 serves dual purposes. First, we developed an optimal reminder task to evaluate whether people could optimally offload intentions. Participants chose between earning maximum points for each remembered item using unaided memory, or earning a smaller amount (which varied from trial to trial) using reminders. This task allowed us to calculate whether people were biased towards using reminders or their own memory, compared with an optimal strategy. If people were biased from optimality, the second level of the purpose was to assess whether such bias could be reduced through metacognitive advice.

In Experiment 2, we used the optimal reminder task to evaluate how optimal older adults were in the cognitive offloading strategies. If older people do not set the same number of reminders as younger adults, the paradigm specifically allowed us to distinguish two possibilities: a) older people used reminders based on their memory abilities such as setting more reminders to compensate for their impaired unaided memory b) older people were more biased towards setting reminders or towards unaided memory regardless of their actual performance.

While Experiment 1 and Experiment 2 used an event-based task to assess intention offloading, Experiment 3 investigated how people chose cognitive offloading strategies using a time-based PM task. Participants were required to press the spacebar at designated times while also performing an ongoing 2-back task. They

could set reminders by clicking a specific button, which means the clock would flash to remind them when the designated time was reached. In the study, we evaluated the effects of three factors, delay length, metacognitive judgement, and accessibility of a clock, on intention offloading.

Experiment 4 used fMRI to investigate whether an instruction to offload information to an external reminder triggers different brain activities to an instruction to forget or remember. To examine the differences, we developed a two-list paradigm inspired by a study of Storm and Stone (2015) where participants were presented with a list of images sequentially followed by an instruction to forget, remember or offload, and then a second list of images. After that, participants were required to answer the position of an image randomly chosen from the two lists (e.g., if the image had been displayed second in the first list, the answer would be “second”). The neural activation of the whole brain during cueing of the instruction was compared across forget, remember and offload trials to reveal the differences.

2

Experiment 1: Optimal Cognitive Offloading and Metacognitive Advice

2.1 Introduction

An effort-based, cost-benefit analysis of the cognitive offloading decision proposes that individuals make decisions by weighing the cost and benefits of the internal memory and external strategies (Risko & Dunn, 2015). It is difficult, however, to quantify the cost and benefit of different strategies with the same measure. For example, there is hardly any uniform way compare the effort of writing a piece of information down with the benefit of not having to worry about forgetting it. In other words, there is no apparent scale to understand the cost and benefit of physical effort and risk avoidance.

To address the lack of unified measure for cost and benefit, in the current study, we developed an optimal reminder task by adapting the intention offloading task (see Chapter 1). In the optimal reminder task, monetary incentives for correct responses depended on the choice of strategy. Participants were offered a choice between earning maximum reward per remembered target when they used internal memory, or a lesser reward per target if they used external reminders. For example, a participant might be offered the choice between using their unaided memory to score 10 points per target, or using reminders to score 5 points per target. If accuracy with reminders was 100% and accuracy without reminders was 40%, it would be optimal to select the option of 5 points with reminders. However, if accuracy without reminders was 60%, it would be optimal to select the option of 10 points without

reminders. The value attached to each target when they set reminders was distributed evenly between zero and the maximum reward (e.g. the maximum reward = 10 points per target, the values = 1-9 points per target) and varied from trial to trial in a random order. Thus, the participants need to weigh their internal strategy with the highest reward per target and the likelihood of less correct targets against the external strategy of reduced reward per target and the possibility of more correct targets. Participants' trade-off between the cost (reduced reward) and benefit (the possibility of better performance) of the offloading strategy would be reflected in a unified measure. Thus, the incentive scheme enables us to evaluate whether the participants chose the optimal strategy or were biased towards one strategy.

Investigating optimality of reminder settings provides possible answers for a theoretical question of whether the biases in strategic offloading decisions could be influenced by metacognitive bias, or the aversion to cognitive effort. Given that previous studies (Boldt & Gilbert, 2019; Gilbert, 2015a, 2015b; see Chapter 1) have shown that people offload intentions based on their metacognitive judgments about their memory abilities, the erroneous metacognitive judgments might lead to a preference for setting reminders, or for using internal memory to complete PM tasks. On the other hand, people were encouraged to choose optimal choices by monetary reward, but greater monetary incentives may not fully offset the effect of the preference to avoid cognitive effort. On the contrary, people might still choose the offloading strategy due to their preference for less cognitive effort even though it means less amount of financial reward for them. It is in line with some studies suggesting that people are willing to get less financial reward to avoid cognitive effort (Apps et al., 2015; Kool et al., 2010; Kurzban et al., 2013; Westbrook et al., 2013). Thus, we divided participants into two groups in the current study. One group

received metacognitive advice when they were free to choose the offloading strategy or unaided memory while the other group did not. The advice about which option was optimal was given on each trial and based on their performance on a previous phase where they were forced to use unaided memory or forced to use the strategy. Consequently, the purposes of the present study are twofold:

- a) Whether people show significant bias from optimality.
- b) If any such bias is reduced when they receive metacognitive advice.

If there is a significant bias, there are two possibilities. One possibility is that the individual bias towards the external or internal strategies is induced by erroneous metacognitive beliefs. If this is the case, providing metacognitive advice could reduce the bias. The other possibility is that people have intrinsic preference to avoid cognitive effort so that they do not always choose the optimal option (Kool et al., 2010; Kurzban et al., 2013), and the bias would not be reduced by metacognitive advice.

2.1.1 Aims

The current study aims to investigate the two hypotheses by carrying out a two-phase experiment investigating the optimality of reminder setting. In the first phase, participants would be forced to use the internal memory or external reminder strategy. By forcing people to use an internal strategy on some trials and an external strategy on the others, it was possible to calculate their accuracy in the two conditions, and hence the optional decision when offered a free choice. In the second phase, participants could choose between the two strategies. In addition, participants were divided into two groups in the second phase of experiment. One group would receive a suggestion in each trial about which choice would be the optimal. Yet, they were also told to choose the strategy freely as they liked. If the bias

in participants' choices was provoked only by metacognitive judgments, the suggestions could rectify the error, thereby diminishing the bias. On the contrary, if the bias was exclusively related to individual preferences such as choosing a suboptimal strategy which reduces cognitive effort, the metacognitive suggestions would have no effect on participants' decisions. This work has been published as Experiment 2 in Gilbert et al. (2020).

2.2 Method and Procedure

2.2.1 Participants

108 paid volunteers (mean age: 31 years, ranges 18-80 years; 37.03 % male; 55 in the advised group and 53 in the unadvised group) who had normal or corrected to normal vision participated in this experiment. A power calculation showing 80% power to detect a between-groups difference with medium effect size ($d = 0.5$) requires at least 102 participants. The effect size was consistent with the effect size which would be expected in a between-groups comparison, where one group exhibited a reminder bias demonstrated by a previous study using the optimal reminder task (Gilbert et al., 2022, Experiment 1), and the other group showed no bias, but had the same standard deviation. One-tailed t test was used because it was hypothesised that metacognitive advice should reduce bias. Participants were randomised into two groups, advised group and unadvised group, with no significant difference in age ($t(95.5) = 1.3, p = .21$) or gender ($\chi^2 = .02, p = .88$). The main goal of this study was to compare the tendency to offload between the two groups. The fact that there were no significant differences in age or gender between the groups indicates that the main finding is unlikely to be influenced by these demographic factors.

2.2.2 Method

See Figure 2.1 for a schematic illustration of the optimal reminder task.

1. Ongoing task: At the beginning of each trial, six yellow circles numbered 1-6 were positioned randomly in a box. Participants were instructed to drag yellow circles in numerical sequence (1, 2, 3, etc.) to the bottom of the box to make them disappear except target circles (mentioned below). A new circle would appear in the same location after one circle disappeared and continued the number sequence (e.g. if number 1-6 was on the screen, after circle 1 was dragged to the bottom of the box, the new circle 7 would appear). This continued until 25 circles had been dragged out of the box.

2. Delayed intention task: Participants were asked to drag yellow standard circles to the bottom edge in the ongoing task. Occasionally, some circles, i.e. target circles, appeared initially in blue, orange or pink and needed to be dragged to the edges in matching colours. To be specific, blue circles should be dragged to the left edge which was also in blue, pink ones to the right edge, and orange ones to the upper edge. The initial colours of the target circles only lasted for 2 seconds and then changed to yellow so that they would look the same with the standard circles when the participants reached them in the sequence. This meant that participants needed to remember the colours of target circles so that they could later drag the circles to respective edges when reaching the circles in the sequence.

3. Offloading strategy: Participants could depend on their own memory to form internal representations of these delayed intentions or offload them on to reminders. They could drag target circles next to the indicated edges as soon as they appeared with particular colours. As a result, the locations of these target circles

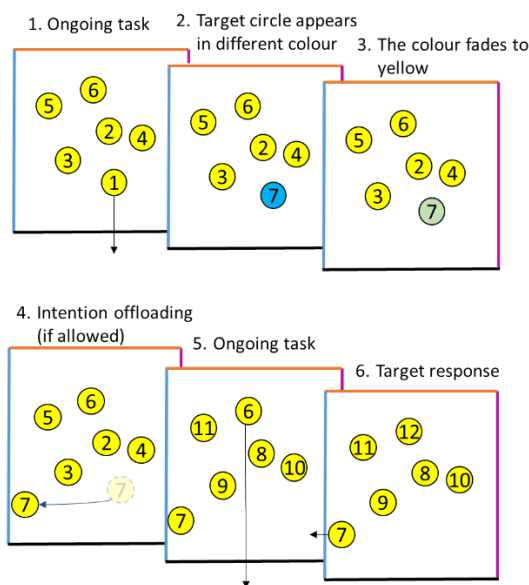
serve as reminders for the delayed intentions.

One trial consisted of a numerical sequence of 1-25. Within this sequence, 10 target circles were allocated to 10 of the numbers from 7-25, which meant the participants would have to remember multiple delayed intentions successively and simultaneously. It seemed impossible for them to remember all of them if the offloading strategy was not allowed. The target circles were allocated randomly to the left, top, and right positions of the box.

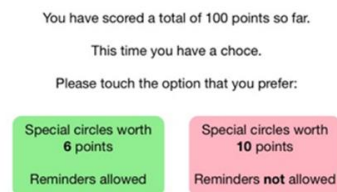
Figure 2.1

Schematic Illustration of the Optimal Reminders Task, and Estimation of Participants' Indifference Points

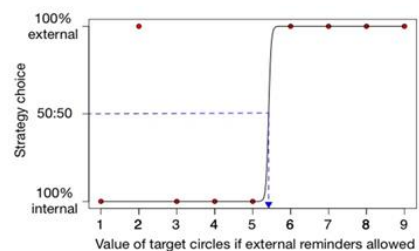
A. Sequence of Events within a Trial



B. Example of Free Choice



C. Actual Indifference Point



2.2.3 Procedure

Participants were tested individually at the Institute of Cognitive Neuroscience, University College London, UK and completed a consent form before the experiment started. The study was approved by the UCL Research Ethics Committee (1584/002).

The optimal reminder task was shown by a tablet (Samsung model SM-T580) with the touchscreen interface. Participants got paid according to the points scored during the task. They would receive £0.3 payment for every 100 points and an extra base payment of £5.

Participants would first have a short practice session of the optimal reminder task. Then, the intention offloading strategy would be explained and they practiced this strategy until they could attain a correct rate of eighty percent. We aimed to ensure that participants had a clear understanding of how to use the strategy. Attaining an accuracy rate of 80% with reminders is not particularly difficult, as most participants are capable of achieving this level on their first or second attempt.

There were two phases in this experiment, compulsory and free-choice. Participants were divided into two groups randomly in the free-choice phase. A timer shown on the screen was set on each trial and participants were encouraged to complete each trial within three minutes. They would be told their total points at the end of each trial.

1. Compulsory Phase: There were eight trials in total. Participants were forced to depend on the internal or external strategy alternatively (with the starting condition counterbalanced between participants). In both cases, they received ten points for each correct target responses. In the forced-internal trials, all circles were immovable in position except the next in the sequence. This meant that participants were forced to use their memory, because the upcoming target circles were fixed in position until it was their turn in the sequence. In the forced-external condition, on the other hand, participants were required to adjust the position of every target circle when they first appeared on the screen, or they could not continue with the

task.

Then, they were asked to report what percentage of target circles they thought they could correctly drag to the instructed location. The following instructions was given: "Now that you have had some practice with the experiment, we would like you to tell us how accurately you can perform the task when you do it without using any reminders. Please use the scale below to indicate what percentage of the special circles you can correctly drag to the instructed side of the square, on average. 100% would mean that you always get every single one correct. 0% would mean that you can never get any of them correct". They responded by dragging a slider on the screen. After they reported their confidence, they were asked: "Now, please tell us how accurately you can perform the task with reminders. As before, 100% would mean that you always get every special circle correct. 0% would mean that you can never get any of them correct."

2. Free-Choice Phase: In the remaining nine trials, subjects were given choices, e.g. scoring 10 points per correct response using their own memory, or 6 points per correct response when allowed to use reminders. These nine trials represented each possible value from 1-9 attached to the target responses with reminders presented in random order. This allows us to measure how optimal people are in the use of reminders. In addition, participants were divided into two groups randomly. While the unadvised group made decisions all by themselves, the advised participants received metacognitive advice indicating which would be the optimal choice to make. However, they were told that they were free to choose either option. These metacognitive suggestions were given by the computer, calculated the optimal strategy based on the performance on the forced-internal and forced-external trials.

At the beginning of the free-choice phase, the participants in the advised group would see the instruction on the screen: "We have been calculating your accuracy on the task so far. This means that we can make a prediction which option is likely to score you most points, based on your performance until now. You will be told this prediction each time you do the task, which may help you to decide whether to do the task with or without reminders. However, you are free to choose whichever option is best - it is completely up to you." On each choice trial they were then given the following information: "According to your performance so far, we have calculated that you will probably score more points if you choose to perform [with/without] reminders. However, you may choose whichever option you prefer." When the expected reward was the same with both strategies, they were told "According to your performance so far, you will score the same number of points regardless of whether you choose to use reminders or not." The unadvised group would not receive the instruction and feedback.

2.2.4 Dependent Measures

1. Self-reported confidence of the internal strategy in forced-internal trials (CI)
2. The average accuracy of target responses in the forced-internal trials (AI)
3. Self-reported confidence of the external strategy in forced-external trials (CE)
4. The average accuracy of target responses in the forced-external trials (AE)
5. Metacognitive bias: the difference between subjective confidence and actual accuracy. This was calculated separately for the internal (i.e., CI – AI) and the external (i.e., CE – AE) condition.

2.2.5 Optimal and actual indifference points:

The optimal indifference point and the actual indifference point were calculated individually to evaluate for each participant whether they have a bias when they choose between two strategies.

A. "The optimal indifference point" (OIP):

The OIP was the value attached to each target circle such that participants would expect the same reward if they chose an external reminder strategy (earning this number of points for each remembered target) or an internal memory strategy (earning 10 points for each remembered target). If participants were offered more than this value, they would earn more points by choosing an external strategy; if they were offered less than this value they would earn more points by using internal memory. For example, if a participant's accuracy is 60% in forced internal trials and 100% in forced external trials, the optimal indifference point will be 6, which means when 6 points are given for each correct response with the external strategy, the participant will get the same points either using the external or internal strategy ($6 \text{ points} \times 10 \text{ correct responses} = 10 \text{ points} \times 6 \text{ correct responses} = 60 \text{ points}$). In line with this, when 8 points are given for each correct response, it is rational to choose the external strategy since the participant might get more points with it ($8 \text{ points} \times 10 \text{ correct responses} = 80 \text{ points}$) than with the internal strategy. On the contrary, when 3 points are given for each correct response with the external strategy ($3 \text{ points} \times 10 \text{ correct responses} = 30 \text{ points}$), the internal strategy would be the more rational choice.

The optimal indifference point can be calculated when we have an average accuracy of each strategy. If participants chose the internal strategy, they always

received 10 points per correct target response. Thus, the expected number of points is $10 \times AI$. The optimal indifference point was the value attached to targets with reminders where the expected number of points would be the same as without. The optimal indifference point is derived below:

$$OIP \times AE = 10 \times AI$$

Rearranging, this gives:

$$OIP = \frac{10 \times AI}{AE}$$

B. "The actual indifference point" (AIP):

Participants actually choose to set reminders when the points of target responses are more than "The actual indifference point" and not when points are less than it. The points given for each correct target responses with the external strategy in free-choice phases were different from trial to trial. The data of the choices were collected across the full range of points from one to nine (Figure 2.1). Then, R package "quickpsy," was used to calculate the actual indifference point by fitting a curve of sigmoid function to the data. At the actual indifference point, participants had 50% probability to choose the external strategy and 50% to rely on their memory. We used this method to obtain the relationship between value and choice of strategy because it can be calculated even if there is not a monotonic relationship between the two.

C. "Reminder bias": the difference between OIP and AIP.

$$Reminder\ bias = OIP - AIP$$

A positive value means the participant tends to set more reminders than would

be optimal. A negative value means the participant tends to set fewer reminders than would be optimal. A participant who had no bias would have a bias score of zero. Importantly, the reminder bias was calculated based on an individual's memory performance. For example, if the AIP is 5, an individual showing 100% accuracy on forced-external trials, and 60% on forced-internal trials, has a reminder bias towards setting reminders. By contrast, if the accuracy on forced-external trials is 100% and on forced-internal trials is 40%, an individual whose AIP is 5 has reminder bias towards unaided memory. Thus, the same AIP could indicate different bias due to the relationship between each individual's optimal strategy and actual choice.

D. "the metacognitive indifference point":

We calculated each participant's indifference point in the same manner as their optimal indifference point but using each participant's self-judged accuracy in the internal and external conditions rather than objective accuracy.

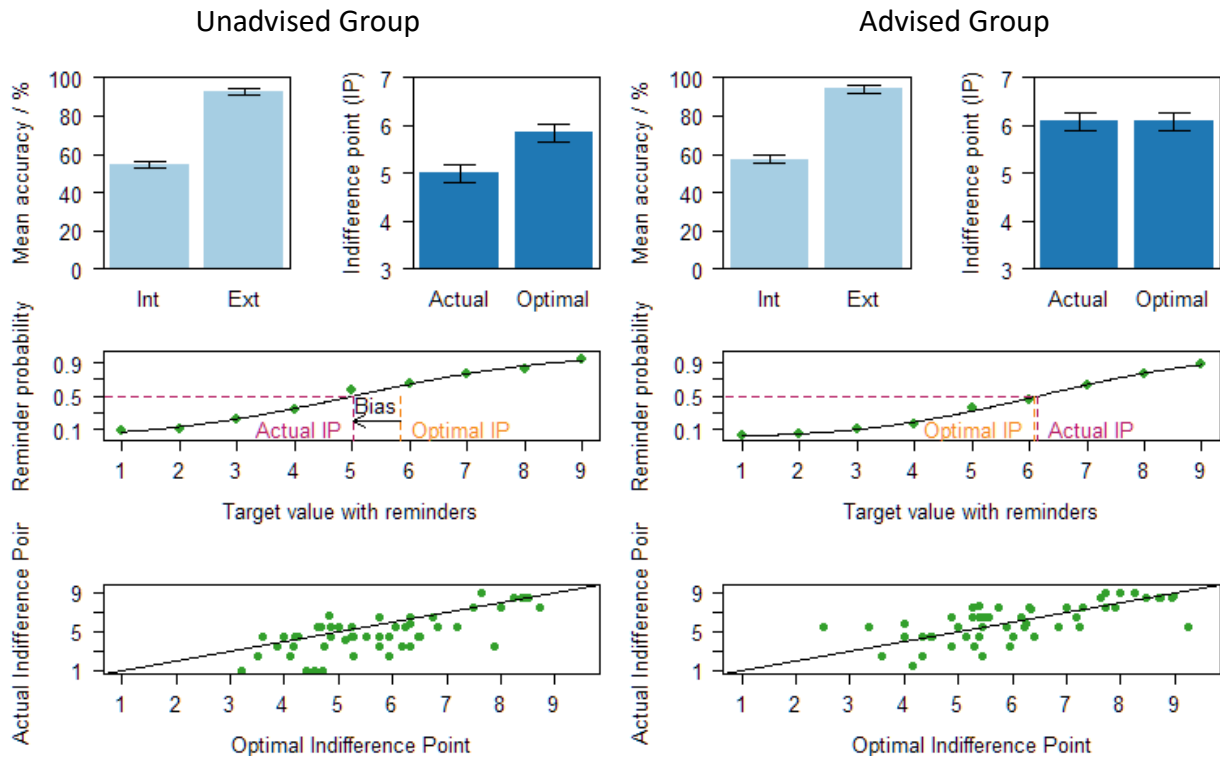
2.3 Results

See Figure 2.2 for a summary of results. The two groups, unadvised and advised, have similar mean trial duration in the forced internal (unadvised group $M=56.6s$, $SD=28.0$; advised group $M=57.0s$, $SD=22.1$) and forced external trials (unadvised group $M=56.3s$, $SD=15.8$; advised group $M=63.7s$, $SD=22.6$). The accuracy in two groups was analysed in a Condition (forced internal vs forced external) x Group (advised, unadvised) ANOVA, showing a main effect of Condition ($F(1, 106) = 786$, $p < .001$, $\eta^2_p = .88$), but no main effect of Group ($F(1, 106) = 1.13$, $p = .290$, $\eta^2_p = .01$) or the Group x Condition interaction ($F(1,106) = .22$, $p = .642$, $\eta^2_p < .01$). Accuracy in the forced internal trials (advised group: $M= 57.5\%$, $SD= 16.5$; unadvised group: $M=$

54.6%, $SD= 16.5$) was much lower than the forced external trials (advised group: $M= 94.0\%$, $SD= 7.2$; unadvised group: $M= 92.4\%$, $SD= 9.0$), which suggests that using reminders improved performance in both groups. The total number of points scored by the advised group ($M= 1210$, $SD= 188$) was higher than the number of points scored by the unadvised group ($M= 1171$, $SD= 203$), however this difference was not statistically significant ($t(104.7)= 1.0$, $p= .30$, $d= .21$).

Figure 2.2

Results from the Optimal Reminder Task: Accuracy and Intention Offloading



Note. Data from the unadvised group are shown on the left and the advised group on the right.

Top row: Light blue: Accuracy in the forced internal (unaided memory) and forced external (reminder) conditions. Error bars represent within-subject confidence intervals that do not overlap with each other, which means $p < .05$. Dark blue: Actual and optimal indifference points.

Middle row: The probability of choosing to set reminders were averaged at every correct target value from 1 to 9 attached to the external strategy in the free-choice trials. The mean of AIP and OIP are also shown.

Bottom row: The AIP and OIP of each participant were shown. The diagonal line represents the calibration that the actual choice is the optimal choice. Points below the line indicate a bias towards external reminders and points above the line indicate a bias towards internal memory.

First, the bias of strategy use in these two groups was evaluated (Figure 2.2). The means of optimal and actual indifference points in the unadvised group were 5.8 and 5.1. The distinction between these two points, a bias towards reminders, was significant ($t(52) = 3.6, p < .001, d = .98$). This implies that participants had a significant bias towards using more reminders than would be optimal. In the advised group, there was no significant decision bias ($t(54) = .3, p = .73, d < .01$). Both the means of optimal and actual indifference points were 6.1. Besides, there was a significant difference of decision bias between these two groups ($t(105) = 2.8, p = .003, d = .55$). In both groups, the optimal and actual indifference points were correlated significantly, which means participants who could get more benefit from the external strategy actually set more reminders. (unadvised: $r(51) = .71, p < .001$; advised: $r(53) = .67, p < .001$).

This brings us to the second point, the metacognitive assessment of the strategies (Figure 2.3). In both groups, the accuracy and confidence were correlated significantly in forced-internal and forced-external trials (unadvised: forced-internal $r(51) = .72, p < .001$; forced-external $r(51) = .81, p < .001$; advised: forced-internal $r(53) = .78, p < .001$; forced-external $r(53) = .59, p < .001$). It is consistent with previous studies (Rummel & Meiser, 2013; Schnitzspahn, Zeintl, et al., 2011; Meeks et al., 2007) that people can predict their actual memory performance to some extent. Additionally, participants were underconfident about the accuracy using internal memory (advised group: $t(54) = 3.6, p < .001, d = .98$; unadvised group: $t(52) = 4.3, p < .001, d = 1.2$). The most important hypothesis about the metacognitive judgments is that the bias towards offloading in the free choice trials should correlate with their underconfidence in the forced internal condition. This hypothesis is supported by finding significant correlation ($r(51) = -.31, p = .026$) between the two in the

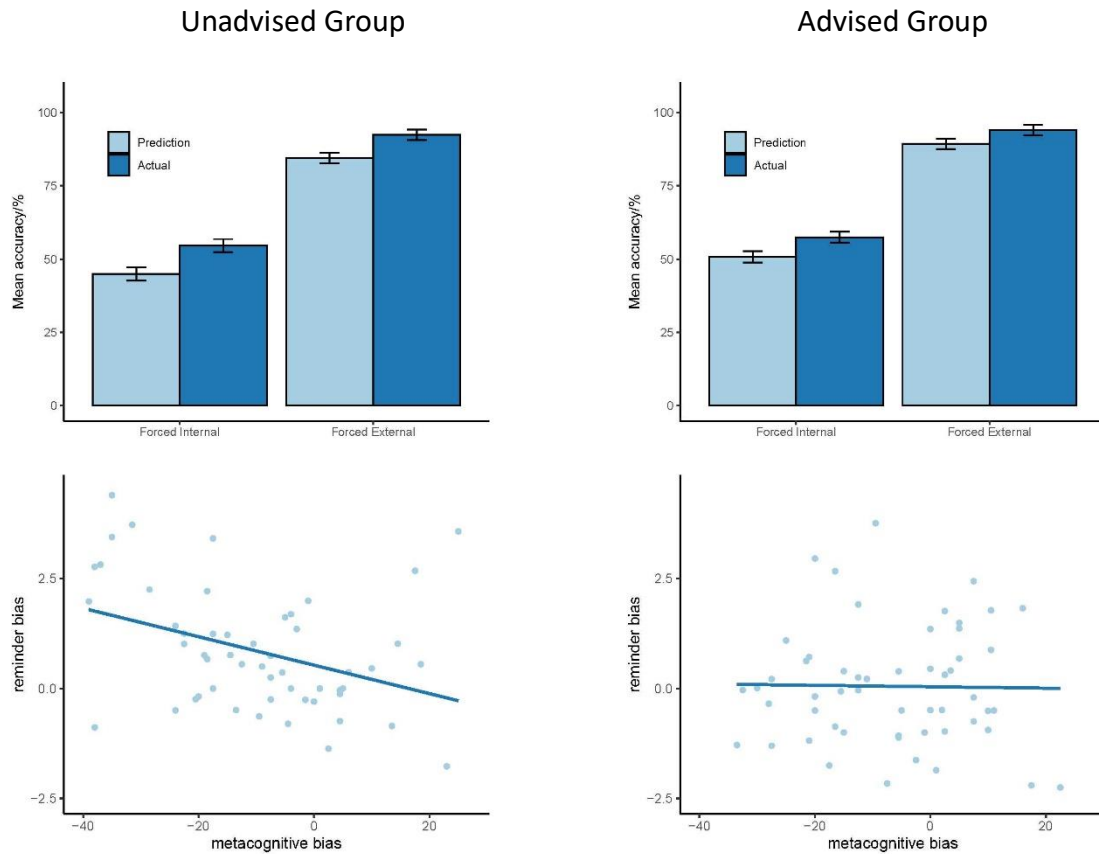
unadvised group. Thus, the extent of underconfidence in internal memory predicted the bias in unadvised participants towards the offloading strategy. Conversely, no significant correlation was observed in the advised group ($r(53) = -.02, p = .89$).

In addition, participants were also underconfident in their performance using external reminders (advised group: $t(54) = 3.40, p = .001, d = .93$; unadvised group: $t(52) = 4.45, p < .001, d = 1.23$). Moreover, for all participants, the metacognitive bias of performance using reminders was correlated with metacognitive bias of performance using internal memory ($r(106) = .32, p < .001$). No significant difference was noted between them by paired-samples t-tests (advised group: $t(54) = -1.11, p = .273, d = .10$; unadvised group: $t(52) = -.64, p = .525, d = .16$). The consistency in these two varieties of metacognitive bias suggests people may have a general underestimation about their abilities to perform this task.

Having shown that the unadvised group had a significant bias towards using more reminders relative to the optimal strategy (i.e., the significant difference in OIP and AIP), we then analysed whether the strategic offloading decisions was different from the optimal option which was calculated based on metacognitive predictions of accuracy in the internal and external conditions rather than objective accuracy. The metacognitive indifference point ($M = 5.33, SD = 2.44$) was not different significantly from the actual indifference point ($M = 5.10, SD = 2.17; t(52) = .74, p = .46, d = .21$). Therefore, the unadvised group's actual choice behaviour was deviated from the optimal strategy based on objective accuracy, but was not deviated from the optimal strategy based on self-judged accuracy.

Figure 2.3

Results from the Optimal Reminder Task: Metacognitive Measures



Note. Data from the unadvised group are shown on the left, and data from the advised group are shown on the right.

Upper row: predicted accuracy versus actual accuracy in the forced internal and forced external conditions (predicted accuracy: Light Blue; actual accuracy: Dark Blue). Error bars representing within-subject confidence intervals do not overlap with each other, which means $p < .05$.

Lower row: The relationship between the reminder bias and metacognitive bias about unaided memory is revealed. While the reminder bias was correlated with the metacognitive bias in the unadvised group, it was not observed in the advised group.

2.4 Discussion

The experiment showed the following findings: a) participants not receiving advice used more reminders than would have been optimal, b) bias was eliminated in the advised group where the participants were given metacognitive advice, and c)

there was significant correlation between objective and subjective indifference point in both the advised and unadvised groups, which means that participants who could get more benefit from the external strategy were more likely to use the external strategy. Two possible explanations can be ruled out for the finding that people were biased towards setting reminders. First, a previous study of Gray et al. (2006) proposed that people choose a strategy because they can make the task less time-consuming, but it cannot explain this finding, seeing as people spent more time when choosing to set reminders. Next, the maximum monetary reward was associated with the optimal strategy so the bias could not be attributed to loss aversion.

On the other hand, at least two potential mechanisms, metacognitive judgment and aversion to cognitive effort, which have been discussed in Chapter 1, could explain the bias towards offloading. To investigate how well participants weigh costs and benefits of the offloading strategy, we allowed participants to choose between the internal memory strategy with the greatest financial incentives per item and the offloading strategy with the possibility of better performance. In spite of a significant correlation between objective and subjective indifference point, the unadvised group had a significant bias towards offloading, i.e. the difference between OIP and AIP. This means that the unadvised group set more reminders than they actually needed even though they were encouraged by financial incentives to choose optimally. By contrast, the bias was eliminated when participants were given metacognitive advice. This finding suggests that the bias towards offloading is possibly ascribed to erroneous metacognitive judgements rather than an intrinsic preference such as avoiding cognitive effort.

2.4.1 Metacognitive Judgement and Reminder Bias

In this study, both groups had basic metacognitive judgements over their ability since subjective confidence correlates with actual accuracy significantly. However, they tended to have underestimated their performance with unaided memory. This underconfidence could predict the degree of the bias towards offloading. These findings are consistent with previous research into PM (Gilbert, 2015b; Meeks et al., 2007; Rummel & Meiser, 2013; Schnitzspahn, Zeintl, et al., 2011) and offer more evidence for the same hypothesis that the bias towards the offloading strategy was provoked by metacognitive bias.

As is discussed above, the present study found that the underconfidence of the offloading strategy was correlated with their underconfidence of unaided memory. Accordingly, these two kinds of metacognitive bias can predict each other although the abilities assessed by each of them are different. This congruity of erroneous metacognitive judgements may be compatible with the suppositions of previous research on domain-general metacognition. Metacognitive judgement in one domain could impact that in another domain. Studies on domain-general metacognition found that the participants' metacognitive confidence in perceptual tasks could be generalized across the perceptual domain and the mnemonic domain since the confidence in the intention offloading and perceptual tasks are strongly correlated (Baird et al., 2013; de Gardelle & Mamassian, 2014; Gilbert, 2015b). Thus, metacognitive confidence might have a general factor which functions across different cognitive activities and people report the confidence as a combination of general confidence and strategy-specific confidence.

2.4.2 Avoiding Cognitive Effort

The result of no reminder bias in the advised group does not fit into an account of an intrinsic bias which avoids internal memory or cognitive effortful processes, seeing as the bias was not found when participants were given advice. However, this result does not show that merely removing metacognitive bias is sufficient to eliminate bias towards offloading in all cases. First, the monetary incentive encouraging participants to optimally use reminders may have contributed to the reduction of bias in the task. Therefore, people could have a bias deviating from the optimal strategy if there is no monetary incentive. Another possible explanation for the result is that participants may simply follow advice to avoid further cognitive load, namely, judge whether to use reminders. It is probable that the reduction in bias in advised group was attributable to the decreased cognitive load resulting from the provision of advice rather than the correction of erroneous metacognitive judgment. Thus, further research is needed to distinguish the effect of cognitive load and metacognitive bias on decisions about reminder usage.

Following the current experiment, a study (Gilbert et al., 2020, Experiment 3) using the optimal reminder task and metacognitive interventions evaluated the influence of metacognition on intention offloading. Instead of being told directly the optimal choice, participants received positive messages such as “ Well done— excellent work! You responded correctly to most of the special circles” or negative messages “Room for improvement. You got some of the special circles wrong.” We manipulated feedbacks which were not deceptive, but had different terms for the same performance, in order to influence metacognition. Also, we manipulated the difficulty of practice trials which influenced metacognitive judgments. Consistent with our results, the study found people’s metacognitive biases were correlated with

reminder biases. Participants receiving both easy practice trials and positive feedback were significantly overconfident in their memory ability whereas participants performing both difficult practice trials and negative feedback were significantly underconfident. The former group had less reminder bias than the latter group, but was still biased towards offloading. Given that overconfidence in the former group did not exhibit a bias against offloading, the reminder bias cannot be fully explained by metacognitive judgment. Moreover, a study of Engeler and Gilbert (2020) found participants receiving metacognitive training, which comprised performance prediction and receiving feedback, showed no metacognitive bias, but remained bias towards offloading. Thus, additional factors could play a role in the decision about reminder settings.

In order to evaluate the additional factors, Sachdeva and Gilbert (2020) manipulated financial incentives to affect reminder bias. Participants receiving performance-dependent rewards had reduced reminder bias than the participants receiving a flat payment. Given the aversion to cognitive load is influenced by monetary reward (Aarts et al., 2010; Padmala & Pessoa, 2011), the findings provided evidence supporting the inference that one of the additional factors is a preference to avoid cognitive effort.

2.4.3 Limitations of this study

The underconfidence of internal memory suggested by the present study may not be observed in the naturalistic settings or real life because of differences in task difficulty. The mean accuracy rate was about 50% when participants depended on their memory in this study, but it is worth noting that most naturalistic tasks may not yield such low performance rates. Thus, in contrast to the results in the laboratory

setting, people may be more confident or sometimes even too confident of their memory to set reminders.

Although the bias towards the offloading strategy provoked by metacognitive bias was observed in the current study, we do not suggest that people being underconfident of their internal memory always set more reminders in all delayed intention tasks in daily life. The metacognitive bias was the foremost factor related to the bias when a monetary incitement and instruction were given, but in everyday life people may choose strategies according to inner bias such as the preference for less cognitive or physical effort that is not offset by the financial reward.

Over- or underconfidence may vary across different populations. Take brain injury patients for example. Knight et al. (2005) assessed the metacognition of healthy adults and those with compromised memory ability due to brain injury. They found that both groups had similar level of confidence. The brain injury group, however, had worse performance, making them overconfident while the control group was underconfident because of better achievement. The results suggest that these patients' metacognitive judgements are not updated after the injury of their memory system. Therefore, people with overconfidence may be inclined to rely on their memory and thus cannot get the full advantage from the external strategies.

2.4.4 Practical implications and future research

The ratio of targets in the PM task in this study was much higher than in common PM paradigms, and the delay in this task was very short. It might raise questions as to whether this paradigm is more akin to dual-tasking than PM. However, a web-based study (Gilbert, 2015a) found that memory performance in this

task could predict the achievement rate of a real-world delayed intention task involving visiting an indicated web link on three specific days in one week.

From our results, several practical implications can be inferred. To begin with, it is not efficient to depend on offloading strategy all the time even though we have increasing opportunities to offload intention onto technological devices with the advancement of modern technology. However, if we offload intentions for all daily activities regardless of their importance and regularity, we may waste our time. For example, some activities such as brushing our teeth are not likely to be forgotten under normal circumstances. The present study evaluated whether people used reminders optimally, and suggests that, people may choose a suboptimal strategy due to erroneous metacognitive judgment. Thus, if the metacognitive bias is rectified, strategic offloading decision can be modified to be more optimal. Learning from past experiences could be one of the ways to ameliorate metacognitive bias. In the current study, the retention time between choosing a strategy and action was short, but in naturalistic tasks, people may wait, after deciding which strategy to use, for a long time before actually accomplishing the intention. Only then can they know how efficient the strategy is. Unfortunately, such a feedback process may be too long for people to establish a strong connection between the strategy and the consequence. As a result, these experiences cannot modify the metacognitive estimation nor drive people to choose an optimal strategy. Consequently, it would be worthwhile to search for innovative and effective approaches to enhance the contribution from experiences.

Those populations with different metacognitive beliefs need specialised education or training for such cognitive rehabilitation (Cicerone et al., 2011) to

reduce their bias, especially because they may be more susceptible to failure in PM tasks. Further research on these populations is needed to evaluate whether improvement of metacognitive bias can help them to choose the optimal strategy since metacognitive control, the other function of metacognition, may deteriorate because of neurological changes.

The results of the current study show that unadvised individuals have a significant bias towards the offloading strategy, against an internal strategy. They also show that this bias can be reduced (or eliminated) by providing metacognitive advice, at least in some circumstances. Therefore metacognitive interventions can potentially help us use external tools more wisely in a world where cognitive function interacts with technology more closely than ever.

Experiment 2: Age Differences in Optimal Cognitive Offloading

3.1 Introduction

PM failures are associated with important real-world implications, such as maintaining health (e.g., forgetting to take medication) and safety (e.g., forgetting to turn off an oven), which may pose a particular challenge for older adults. Moreover, it has been shown that PM failures have greater impact than retrospective memory on the functional independence of older adults (Hering et al., 2018; Sheppard et al., 2020), and yet forgetting to perform an intended action at the appropriate future moment is reported as the most frequent memory failure in our everyday life (Haas et al., 2020). External reminders such as calendars, alarms, and digital devices can play an important role in reducing PM failures in everyday life (Jones et al., 2021), however such strategies are not always chosen optimally when one considers that setting a reminder, as well as increasing the likelihood of remembering, also carries a cost in terms of time and effort (Gilbert et al., 2020). Thus, in Experiment 2, we aimed to directly compare how optimal younger and older adults are when they make these choices in an experimental task. By understanding age-related changes in reminder usage, this could lead to the development of interventions to improve the fulfilment of PM tasks in older adults.

3.1.1 Prospective Memory and Aging

Contrary to retrospective memory tasks in the laboratory, where the

experimenter typically initiates retrieval, PM tasks pose a high demand on self-initiated processes and offer low environmental support (Craik, 1986). Since the ability to recruit self-initiated processes declines with advancing age, it has been suggested that PM tasks should be particularly sensitive to the effects of ageing (Maylor, 1995; McDaniel & Einstein, 2000). Indeed, in laboratory settings, younger participants usually outperform older participants on PM tasks (Henry et al., 2004; Kliegel, Jäger, et al., 2008; Zuber & Kliegel, 2020). However, age differences in PM varies substantially across individual studies. Whereas some studies found a significant age-related decline in PM (e.g., Park et al., 1997), other revealed similar PM performance in younger and older adults (e.g., Einstein & McDaniel, 1990). Although the underlying mechanisms are still not fully understood, recent studies have identified some of the factors that affect age-related PM, which can be broadly categorised into cognitive, task-inherent, and context factors.

Among the cognitive factors that accentuate age differences in PM, executive functions have received considerable attention in the literature (Zuber & Kliegel, 2020). Previous research has shown that older adults tend to exhibit a particularly marked decline in PM tasks that place a high demand on working memory (Bisiacchi et al., 2008; Logie et al., 2004). This suggests that the ability to maintain information in working memory plays a crucial role in the successful accomplishment of PM tasks. Furthermore, inhibition – i.e., the ability to refrain from performing a particular action, or to ignore distracting information – and shifting – i.e., the capacity to switch attention from one task to another – also tend to decline with older age and may account for age differences in PM (Azzopardi et al., 2017; Schnitzspahn et al., 2013).

Among the task-inherent factors associated with age differences in PM, an

important role is played by the attentional demands of a PM task. In this context, the 'focality' of a task is a key concept. A focal task can be defined as one where an individual already attends, as part of an ongoing task, to the stimulus feature that cues a delayed intention. According to McDaniel & Einstein (2000), such tasks can potentially be achieved by spontaneous retrieval of the intended action at the appropriate time. By contrast, nonfocal tasks require an individual to attend to a different stimulus feature in order to detect the PM cue. This type of task is hypothesised to require deliberate controlled attention towards the cue-defining feature (McDaniel & Einstein, 2000). Under the general assumption that older adults have reduced ability for controlled attention, age differences are expected particularly for nonfocal tasks (McDaniel & Einstein, 2000; Rose et al., 2010b). The results of a meta-analysis conducted by Kliegel, Jäger, et al. (2008) confirmed this prediction. Conversely, PM age differences are reduced when attention is directed towards the PM task, such as under low ongoing task absorption (Schnitzspahn, Ihle, et al., 2011), or when the importance of the PM task is higher or emphasised (Hering et al., 2014).

In addition to these cognitive and task-inherent factors, age differences in PM are further influenced by the setting in which the prospective task must be performed. That is, the age-related deficit often documented in laboratory-based studies is typically reduced or even reversed in naturalistic studies, where the PM task is embedded in the participants' everyday lives, for example by asking them to make telephone calls or send postcards over a period of days or weeks (Henry et al., 2004). This pattern, including an age-related deficit in laboratory tasks and an age-related benefit in naturalistic tasks, has been described as "paradoxical" (Rendell & Craik, 2000; Rendell & Thomson, 1999). Although some research suggested that this

age PM paradox could be explained by the greater use of reminders in naturalistic settings (Craik & Kerr, 1996; Dobbs & Reeves, 1996; Glisky, 1996), this explanation has not been supported by more recent investigations (Aberle et al., 2010; Phillips et al., 2008). For example, Ihle and colleagues (2012) showed that although reminder usage was positively associated with PM performance, it did not eliminate (but even increased) the age effect.

3.1.2 Metacognition and Aging

Given the critical contribution of metacognition to the decision about setting reminders, it is worth examining the aging processes of the metacognitive functioning. There is considerable evidence that brain regions in prefrontal and parietal cortex which are important for metacognition (Fleming, Huijgen, et al., 2012; McCurdy et al., 2013) are also predisposed to ageing-related atrophy (Tisserand et al., 2004). However, behavioural evidence for metacognitive decline in older age is mixed. While some elements of metacognitive functioning may decline with age, other elements may be preserved or even improved (see Castel et al., 2016; Hertzog, 2016, for an overview). This may depend on the type of task, memory domain, assessment method, and so on (McGillivray & Castel, 2017; Siegel & Castel, 2019).

For example, studies investigating judgements of learning (JOLs) have produced mixed results. Bruce et al. (1982) investigated differences in metacognitive monitoring between younger and older people by asking participants to predict how many items they would remember after learning a word list. Younger adults demonstrated greater prediction accuracy (i.e., a smaller difference between the predicted and actual number of recalled words). The two groups predicted a similar number of words, but older participants recalled fewer. As a result, older people

exhibited overconfidence in their memory performance, which could be interpreted as an age-related metacognitive deficit. However, more recent studies investigating JOLs on the basis of item-by-item predictions have suggested preserved metacognitive monitoring in older participants. Connor, Dunlosky, and Hertzog (1997) asked participants, after studying each item, to predict whether they would remember it. They found comparable prediction accuracy between younger and older participants. Studies of item-level JOLs have also shown comparable ability of younger and older participants to distinguish remembered versus forgotten items (Devolder et al., 1990; Hertzog et al., 2010).

Also within the domain of PM, the relatively few studies of metacognition and ageing have produced mixed results (Kuhlmann, 2019). Devolder et al. (1990) asked participants to predict the likelihood of performing a naturalistic PM task and found older adults displayed better metacognitive judgement. Cauvin et al. (2019) used a laboratory paradigm to evaluate age differences in metacognitive functioning for two components: prospective versus retrospective. Having studied word pairs representing cue-action associations, participants were asked to press a specific button when they noticed one of the cues, then type its associated pair. They provided separate predictions for how likely they would be to press the button (i.e., the prospective component) and how likely they would be to remember the paired word (i.e., the retrospective component). While there were no age differences for the retrospective component, older adults were more overconfident of their performance for the prospective component, which was considered as an age-related deficit. These results suggest that ageing has a greater impact on metacognitive monitoring for the prospective component than the retrospective component of PM. This dissociation highlights the importance of distinguishing

different aspects of metacognition, which may not be influenced uniformly by ageing.

In contrast with the evidence above suggesting that older adults may be more overconfident about their memory ability than younger adults, in a recent study, Scarampi and Kliegel (2021) found that both age groups were similarly underconfident when asked to predict their performance at a PM task (Scarampi & Kliegel, 2021). Still, other studies have shown that confidence in one's memory ability tends to decline with age (Dobbs & Rule, 1987). As a result, it might be expected that older adults would be more likely to use external reminders. Consistent with this, it has been suggested that lower confidence in older adults can lead to a volitional avoidance of memory retrieval (Touron, 2015). However, studies investigating self-report of reminder usage in different age groups have produced inconsistent results (Lovelace & Twohig, 1990; Rendell & Thomson, 1999).

3.1.2.1 Metacognitive control and aging

The age difference in metacognitive control was revealed in a study of Redshaw et al.(2018) where children were asked to fulfil one or three delayed intentions in each trial and allowed to choose freely between using internal memory and setting reminders in the intention offloading task (Gilbert, 2015a). Only older children (older than 9 years old) set more reminders in trials of three delayed intentions, which required more cognitive demand. Given younger children (< 9 years) had equivalently accurate metacognitive judgment, the phenomenon might result from less engagement in translating the outcome of the metacognitive monitor into the strategical decision about setting reminders.

A small number of studies have so far investigated whether older adults would

offset age-related declines in memory ability by setting more reminders (Einstein & McDaniel, 1990a; Henry et al., 2012). The findings generally indicated that although the availability of reminders enhanced PM performance, this effect did not interact with age. These results were confirmed by a recent study by Scarampi and Gilbert (2021) employing the intention offloading task (Gilbert, 2015a). Older adults had poorer unaided performance. When they were permitted to set reminders, they were only slightly (and non-significantly) more likely to do so, and the performance gap between younger and older adults remained. Therefore, older did not fully compensate for reduced memory capacity by using reminders. In addition, older (but not younger) adults were overconfident about their memory ability. Based on this, Scarampi and Gilbert (2021) concluded that older adults do not necessarily compensate for impaired memory ability and that metacognitive differences between younger and older individuals may account for this, at least in part.

The aim of this study was to investigate whether two age groups would use reminders optimally and the relationship between reminder-setting and metacognitive evaluations in two age groups. Previous two intention offloading studies showed mixed results for aging effects. While Gilbert (2015a) showed that older adults tend to set more reminders than younger adults accomplishing a PM task, Scarampi and Gilbert (2021) suggested older adults do not fully compensate impaired memory ability by setting more reminders. Therefore we used the optimal reminder task which could evaluate two aspects of aging effects on intention offloading: a) the absolute number of reminders, and b) the pro- or anti- reminder bias (bias towards or against setting reminders), relative to the optimal strategy. If older adults set more reminders, this could reflect a) an adaptive compensatory response to a reduced unaided ability to perform the task; b) a shift in the bias

towards using, or avoiding external reminders; or c) a combination of the two.

3.1.3 Hypothesis

We would first predict that the older population might show the greater use of external reminders based on the previous two studies (Gilbert 2015b, Scarampi & Gilbert, 2021). What about the issue of pro- versus anti-reminder bias? The evidence reviewed above suggests that older adults can be over-confident in their PM ability. However, the wider cognitive ageing literature is mixed and also provides substantial evidence that older adults tend to avoid internal memory processes (Touron, 2015), and are more reliant on environmental support (Lindenberger & Mayr, 2013) or scaffolding (Zahodne & Reuter-Lorenz, 2019). Based on this evidence, we initially hypothesised that older adults may show an increased bias towards external reminders, along with an increased propensity to set reminders in general. As shown below, these hypotheses were only partially correct. Before data collection, we pre-registered our hypotheses, experimental procedure, and analysis plan (<https://osf.io/gmrbe/>).

3.2 Method and Procedure

3.2.1 Participants

The aim was to include a total of 44 younger (age between 18 and 30 years old) and 44 older (age more than 60 years old) participants, according to our pre-registered plan. This was required in order to achieve 80% power to detect an effect size (Cohen's d) of 0.61, in a two-tailed test with an alpha of 0.05. This effect size was based on the meta-analysis of Uttl (2008), in the comparison between younger and older participants in the performance of PM paradigms most similar to the present

one (“vigilance and event-based tasks”). In order to achieve this sample size, a total of 109 volunteers were tested due to our pre-registered exclusion criteria (see below).

Younger participants were recruited from the Institute of Cognitive Neuroscience participant database, while older participants were recruited via flyers distributed by email and around the university campus and local community. Participants provided brief health histories to allow researchers to check inclusion criteria before they were invited to attend. They were excluded if they had history of major neurological or psychological conditions, significant mental or memory problems diagnosed by a doctor, or colour-blindness. All participants provided informed consent before taking part and the study was approved by the University College London (UCL) Research Ethics Committee (1584/002).

3.2.2 Method

See Figure 3.1 for a schematic illustration of the optimal reminder task.

1. Ongoing task: At the beginning of each trial, six yellow circles numbered 1-6 were positioned randomly in a box. Participants were instructed to drag yellow circles in numerical sequence (1, 2, 3, etc.) to the bottom of the box to make them disappear except target circles (mentioned below). A new circle would appear in the same location after one circle disappeared and continued the number sequence (e.g. if number 1-6 was on the screen, after circle 1 was dragged to the bottom of the box, the new circle 7 would appear). This continued until 25 circles had been dragged out of the box.
2. Delayed intention task: When new circles appeared on the screen, occasionally they were presented initially in blue, orange, or pink. This served as an instruction for

a delayed intention that the circle should be dragged to the blue (left), pink (right), or orange (top) edge when they were reached in the sequence. For example, if number 7 initially appeared in blue, the participant should drag numbers 2-6 to the bottom of the box, then drag number 7 to the left. The initial colours of these target circles only lasted for 2 seconds, then faded to yellow so that they were identical to the other circles. Therefore, participants needed to remember the colours of target circles and drag them to their respective edges later.

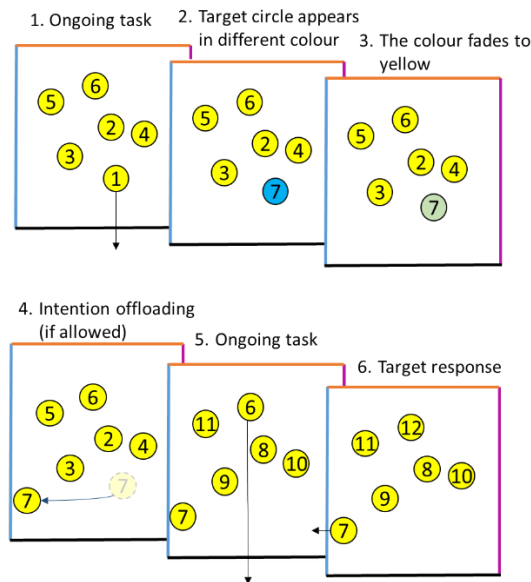
3. Offloading strategy: Participants could depend on their own memory to form internal representations of the delayed intentions or offload them by setting external reminders. They did this by dragging target circles next to the intended edge of the box as soon as they first appeared. This meant that the location of the target circles served as a reminder for where it should eventually be dragged when it was reached in the sequence.

Within each 25-circle sequence, 10 targets were allocated to the number 7-25, spaced as evenly as possible. This means that participants would need to remember multiple delayed intentions simultaneously and it was unlikely that they would remember all of them if the offloading strategy was not allowed. The target circles were allocated randomly to the left, top, and right positions of the box.

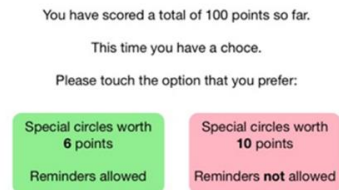
Figure 3.1

Schematic Illustration of the Optimal Reminders Task, and Estimation of Participants' Indifference Points

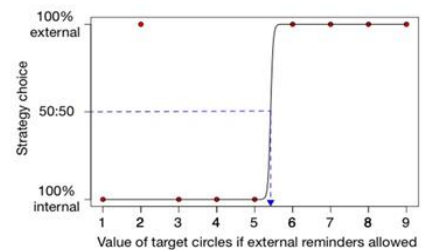
A. Sequence of Events within a Trial



B. Example of Free Choice



C. Actual Indifference Point



3.2.3 Procedure

Participants were tested individually at the Institute of Cognitive Neuroscience, UCL. They performed the experimental task using the touchscreen of a tablet computer (Samsung SM-T580). Participants were paid according to the points they scored during the task. They received £0.30 for every 100 points, along with a base payment of £8.50 so that the maximum reward was about £10.

Participants first had a short practice of the optimal reminders task relying on their internal memory only. Then the intention offloading strategy was explained and they practiced this strategy until they achieved an 80% accuracy rate. After this, participants made a metacognitive judgement to indicate what percentage of target circles they thought they were able to remember, separately for when they had to use their own memory and external reminders. The instructions were: "Now that you

have had some practice with the experiment, we would like you to tell us how accurately you can perform the task when you do it without using any reminders. Please use the scale below to indicate what percentage of the special circles you can correctly drag to the instructed side of the square, on average. 100% would mean that you always get every single one correct. 0% would mean that you can never get any of them correct". After participants reported their confidence, they were asked "Now, please tell us how accurately you can perform the task with reminders. As before, 100% would mean that you always get every special circle correct. 0% would mean that you can never get any of them correct".

There were 17 trials in total (each consisting of 25 circles to be dragged, including 10 targets). For the eight even-numbered trials, participants were forced either to use their own memory ("forced internal" condition) or to use reminders ("forced external condition").

They alternated between these conditions, with the starting condition randomised. In these trials, they got ten points for each correct target response. In forced-internal trials, all circles were immovable in position except the current item in the sequence, so target circles could not be dragged into reminder locations. In the forced-external trials, participants were required to adjust the position of each new target circle or they were not able to continue with the task.

For the nine odd-numbered trials, participants were given a free choice between earning maximum points (10) for each remembered target using their own memory, or a smaller number of points using reminders. The smaller number varied from trial to trial, with the nine possible values from 1-9 presented in random order. During each trial, a timer on the screen counted down from three minutes and

participants were encouraged to complete each trial before it reached zero. Following each trial, participants were told the total number of points they had scored so far. For a full demonstration of the experimental task (including the full practice session), please see:

<http://samgilbert.net/demos/optimalDemo/start.html>

After the optimal reminder task, all participants were administered some further tests in the following order: Beck's Depression Inventory (BDI), Montreal Cognitive Assessment Test (MoCA), National Adult Reading Test (NART) and Raven's Progressive Matrices – form A (RPM). BDI and MoCA were adopted to exclude participants with depression or suspected dementia. The cut-off points of the BDI and MoCA were 11 (Suija et al., 2012) and 23 (Luis et al., 2009) respectively. The NART was used to measure crystallised intelligence and the RPM to investigate fluid intelligence (Bilker et al., 2012).

3.2.4 Dependent Measures

1. Self-reported confidence of the internal strategy in forced-internal trials (CI)
2. The average accuracy of target responses in the forced-internal trials (AI)
3. Self-reported confidence of the external strategy in forced-external trials (CE)
4. The average accuracy of target responses in the forced-external trials (AE)
5. Metacognitive bias: the difference between subjective confidence and actual accuracy. This was calculated separately for the internal (i.e., CI – AI) and the external (i.e., CE – AE) condition.

6. Optimal indifference point (OIP): the target value at which an unbiased individual would be indifferent between using internal memory (earning 10 points per remembered item) or external reminders (earning this number of points per remembered item).

For example, if a participant's accuracy was 60% in forced internal trials and 100% in forced external trials, the optimal indifference point would be 6 because the total number of points scored in the internal condition (60% accuracy x 10 points per item) is the same as the external condition (100% accuracy x 6 points per item). Seeing as targets are always worth 10 points in the internal condition, we can derive:

$$OIP \times AE = 10 \times AI$$

Rearranging, this gives:

$$OIP = \frac{10 \times AI}{AE}$$

7. "The actual indifference point" (AIP): the estimated point at which participants were actually indifferent between the two strategies. If this is higher than the OIP, this indicates a bias towards internal memory (because participants would need to be offered a greater amount than the OIP to choose external reminders). If it is lower than the OIP, this indicates a bias towards external reminders (because participants would be using external reminders even when offered a number of points below the OIP). The AIP was calculated by fitting a sigmoid function to the choice data across the 9 trials using the R package "quickpsy" bounded to the range 1-9 and otherwise using default parameters. This allowed us to calculate the value associated with a 50% probability of

choosing either strategy, according to this function. This approach does not necessarily require a monotonic relationship between value and strategy choice, for example, if participants accidentally chose an external strategy for one of the low-value choices (see Figure 3.2). The AIP can be taken as an index of each individual's propensity to use external reminders (low AIP = high number of reminders, and vice versa).

8. "Reminder bias": the difference between the OIP and AIP (i.e., OIP minus AIP). A positive value would indicate that the participant set more reminders than would be optimal. A negative value would mean that the participant set fewer reminders than optimal. A participant who had no bias had a score of zero. Note that the reminder bias depends on an individual's own level of memory performance. For example, an AIP of 5 would indicate a bias towards internal memory for an individual who achieves 40% accuracy using internal memory, assuming that accuracy is 100% with reminders. The same AIP would indicate a bias towards external reminders for an individual who achieves 60% accuracy with internal memory. Therefore, this bias score is relative to each individual's optimal strategy. It is not the same as the overall propensity to set reminders.

3.2.5 Exclusion Criteria

Participants were excluded if they satisfied any of the following pre-registered criteria. The exclusion criteria were consistent with those used in a previous study by Gilbert and colleagues which utilised the same task (Gilbert et al., 2020, Experiment 3):

- The cut-off points of the BDI and MoCA were 11 (Suija et al., 2012) and 23 (Luis

et al., 2009) respectively.

- Accuracy in the forced internal condition lower than 10%. This excluded participants who did not concentrate on the task during the experiment.

- Accuracy in the forced external condition lower than 70%. This excluded participants who did not fully realise how to use the strategy.

- Negative point biserial correlation between points offered for correct responses on each trial using reminders (1-9) and choice of strategy (0=own memory, 1=reminders). This excluded participants if they were more likely to set reminders when it earned them fewer points, which would suggest random strategy selection behaviour

- Reminder bias score more than 2.5 standard deviations from the mean of participants in that age group

- Metacognitive bias with unaided memory more than 2.5 standard deviations from the mean of participants in that age group

Twenty-one participants were excluded due to our pre-registered criteria (<https://osf.io/gmrbe/>). Seventeen (5 younger; 12 older) were removed due to BDI or MoCA scores. Two (1 younger; 1 older) were removed as a result of the negative point biserial correlation. Two younger participants were excluded because their reminder bias score was more than 2.5 standard deviations from the mean. No participant met any of the other exclusion criteria.

3.3 Results

A total of 109 volunteers to reach the sample of 88 participants were tested. (Younger: 44, mean age: 23.82 years, ranges 19-30 years, male 13; Older: 44, mean age: 72.79 years, ranges 60-89 years, male 16), was analysed statistically.

See Figure 3.2, and Figure 3.3 for a summary of results. All analyses were conducted in accordance with the pre-registered plan, except where noted. There was no significant difference in education duration between younger ($M = 16.9$; $SD = 2.6$) and older ($M = 16.4$, $SD = 4.9$) participants ($t(66.4) = 0.68$, $p = .50$). RPM scores were higher in younger ($M = 6.3$, $SD = 2.5$) than older ($M = 4.6$, $SD = 1.7$) participants ($t(76.5) = 3.7$, $p < .001$, $d = 0.78$). By contrast, NART scores were higher in older ($M = 37.6$, $SD = 6.5$) than younger ($M = 27.1$, $SD = 5.1$) participants ($t(81) = 8.40$, $p < .001$, $d = 1.80$). These results are consistent with previous research suggesting that healthy older people maintain crystallised intelligence while fluid intelligence tends to decline (Horn & Cattell, 1967; Nettelbeck & Rabbitt, 1992).

3.3.1 Accuracy at the Delayed-intention Task

We first investigated accuracy in the forced-internal and forced-external conditions (Figure 3.2). Accuracy in the forced-internal condition was considerably higher in younger ($M = 66.4\%$, $SD = 17.3$) than older ($M = 42.5\%$, $SD = 11.6$) participants ($t(76.9) = 7.90$, $p < .001$, $d = 1.67$). The forced-external condition also showed a small but statistically significant advantage in the younger ($M = 98\%$, $SD = 3.4$) compared with the older ($M = 95.0\%$, $SD = 5.7$) participants ($t(61.9) = 3.60$, $p < .001$, $d = 0.80$).

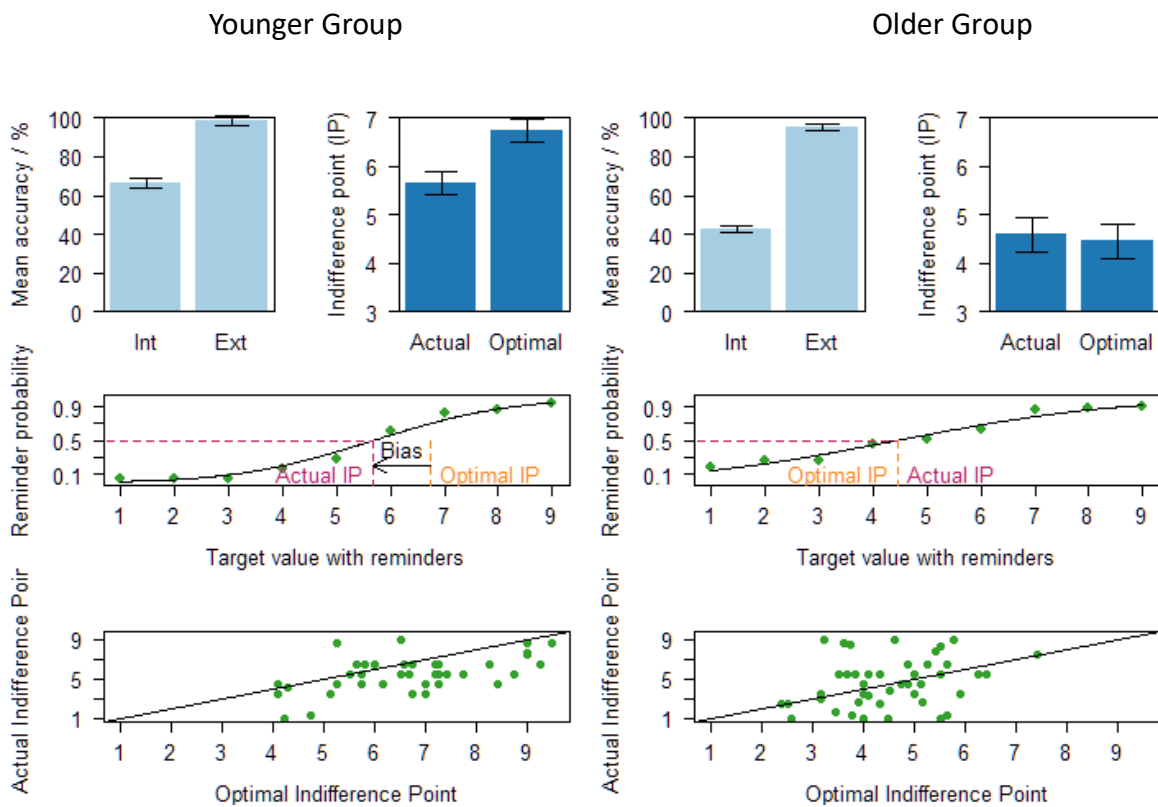
3.3.2 Reminder Setting Behaviour

Next, we investigated the total number of trials, out of the 9 choice trials, where participants chose to use external reminders. This number was significantly higher in the older ($M = 5$, $SD = 2.2$) than the younger ($M = 3.8$, $SD = 1.8$) group ($t(82.2) = 2.67$, $p = .009$, $d = 0.57$). Relatedly, the AIP was significantly lower in the older ($M = 4.6$, $SD = 2.4$) than the younger ($M = 5.6$, $SD = 1.8$) group ($t(78.5) = 2.32$, $p = .02$, $d = 0.49$).

Therefore, older adults used more reminders than the younger adults. The OIP was also significantly lower in the older ($M = 4.5, SD = 1.1$) than the younger ($M = 6.7, SD = 1.6$) group ($t(77.3) = 7.64, p < .001, d = 1.63$). This shows that it was optimal for older adults to use more reminders than younger adults.

Figure 3.2

Results from the Optimal Reminder Task: Accuracy and Intention Offloading



Note. Data from the younger group are shown on the left and the older group on the right.

Top row: Light blue: Accuracy in the forced internal (unaided memory) and forced external (reminder) conditions. Error bars represent within-subject confidence intervals that do not overlap with each other, which means $p < .05$. Dark blue: Actual and optimal indifference points.

Middle row: The probability of choosing to set reminders were averaged at every correct target value from 1 to 9 attached to the external strategy in the free-choice trials. The mean of AIP and OIP are also shown.

Bottom row: The AIP and OIP of each participant were shown. The diagonal line

represents the calibration that the actual choice is the optimal choice. Points below the line indicate a bias towards external reminders and points above the line indicate a bias towards internal memory.

3.3.3 Reminder Bias

Having shown 1) that it was optimal for older adults to use more reminders than younger adults, and 2) that they actually did so, we next investigated the reminder bias, i.e., the difference between actual and optimal reminder-setting behaviour. The younger group showed a significant pro-reminder bias ($M = 1.1$, $SD = 1.6$; $t(43) = 4.53$, $p < .001$, $d = 0.68$), however there was no significant bias in the older group ($M = -0.1$, $SD = 2.3$; $t(43) = 0.37$, $p = .72$, $d = 0.06$). Moreover, the pro-reminder bias was significantly greater in the younger than the older group ($t(75.6) = 2.84$, $p = .006$, $d = 0.61$). These results are not congruent with our initial hypothesis. Rather than an increased pro-reminder bias in older adults, as we initially predicted, the pro-reminder bias was actually eliminated in the older group.

3.3.4 Metacognitive Judgements

Older adults predicted lower accuracy with internal memory ($M = 36.4$, $SD = 13.6$) than the younger group ($M = 46.1$, $SD = 17.5$; $t(81.0) = 2.90$, $p = .005$, $d = 0.62$). Both groups were significantly underconfident, relative to their actual accuracy level (younger: $M = -20.4$, $SD = 21.3$, $t(43) = 6.4$, $p < .001$, $d = 0.96$; older: $M = -6.1$, $SD = 17.0$, $t(43) = 2.37$, $p = .02$, $d = 0.36$). The degree of underconfidence was greater in younger than older participants ($t(82.1) = 3.47$, $p < .001$, $d = 0.74$).

Older adults also predicted lower accuracy when using external reminders ($M = 74.6$, $SD = 17.5$) than the younger group ($M = 92.0$, $SD = 9.4$; $t(65.9) = 5.9$, $p < .001$, $d = 1.25$). Again, both groups were significantly underconfident (younger: $M = -6.4$, SD

= 9.8; $t(43) = 4.3, p < .001, d = 0.65$; older: $M = -20.5, SD = 17.1; t(43) = 7.92, p < .001, d = 1.19$). However, this time the degree of underconfidence was greater in older than younger participants ($t(68.4) = 4.73, p < .001, d = 1.01$).

We investigated the correlation between internal metacognitive bias and reminder bias separately in the two groups. The expected negative correlation was obtained in younger individuals ($r = -.48, p < .001$), showing that participants who were more underconfident in their memory ability tended to exhibit a greater pro-reminder bias. However, the correlation was not significant in the older group ($r = -.21, p = .17$).

3.3.5 Additional, Non-preregistered Analyses

This section reports some exploratory tests conducted in addition to the pre-registered ones described above. First, we investigated the correlation between AIP and OIP separately in the two groups. Both correlations were significant (younger: $r = .56, p < .001$; older: $r = .32, p = .03$). This shows that individuals who had a greater need for reminders (lower OIP) also tended to set them more often (lower AIP). This suggests that individuals in both groups used metacognitive judgements to influence their reminder-setting behaviour.

Second, we investigated accuracy in the forced-internal and forced-external condition in a mixed 2x2 ANOVA with factors Age and Condition (the significant effects of Age, separately for each condition, are already reported above). There were significant main effects of Age ($F(1,86) = 55.72, p < .001, \eta_p^2 = .39$) and Condition ($F(1,86) = 809, p < .001, \eta_p^2 = .9$) along with a significant interaction ($F(1,86) = 50.4, p < .001, \eta_p^2 = .37$). This interaction shows that the age-related

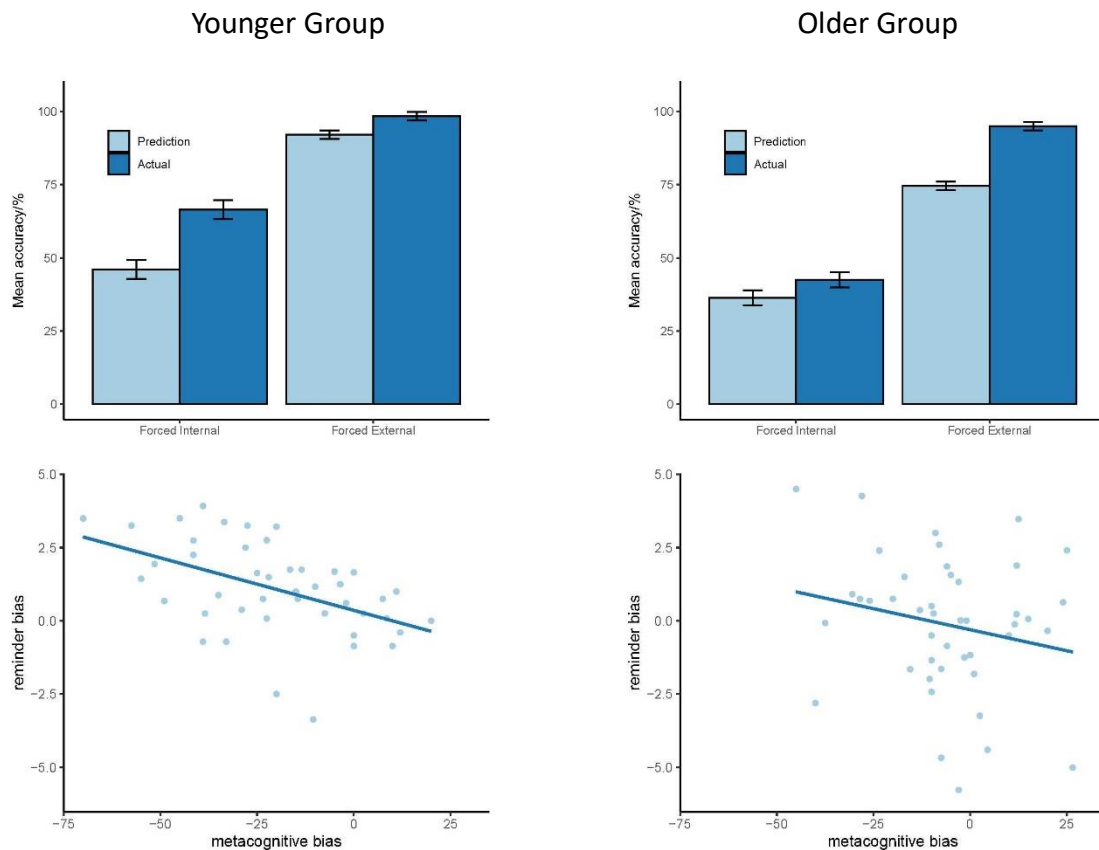
impairment found in the forced-internal condition was significantly attenuated when reminders were used in the forced-external condition.

Third, we directly compared the degree of underconfidence in the forced-internal versus forced-external conditions, along with the predicted versus actual benefit of reminders. The younger group were significantly more underconfident about the forced-internal than the forced-external condition ($t(43) = 4.04, p < .001, d = 0.61$). Consistent with this, the predicted benefit of reminders in younger participants ($M = 46.0\%, SD = 18.3$) was larger than the actual benefit ($M = 31.5\% SD = 16.4; t(43) = 3.80, p < .001, d = 0.58$). The older group showed the reverse pattern. They were significantly more underconfident in the forced-external than the forced-internal condition ($t(43) = 4.22, p < .001, d = 0.64$). Consistent with this, the predicted benefit of reminders in older participants ($M = 38.1\%, SD = 19.7$) was less than the actual benefit ($M = 52.5\% SD = 10.8; t(43) = 4.22, p < .001, d = 0.64$).

Finally, seeing as there was a significant correlation between metacognitive bias and reminder bias in the younger but not the older group, we directly compared these correlations using a Fisher r -to- z transformation. The result showed that there was no significant group difference ($z(41) = 1.4, p = .16$).

Figure 3.3

Results from the Optimal Reminder Task: Metacognitive Measures



Note. Data from the younger group are shown on the left, and data from the older group are shown on the right.

Upper row: predicted accuracy versus actual accuracy in the forced internal and forced external (predicted accuracy: Light Blue; actual accuracy: Dark Blue) conditions. Error bars represent within-subject confidence intervals do not overlap with each other, which means $p < .05$.

Lower row: The relationship between the reminder bias and metacognitive bias about unaided memory is revealed. While the reminder bias was correlated with the metacognitive error in the younger group, it was not observed in the older group.

3.4 Discussion

This study was designed to investigate how ageing influences decisions about the use of offloading strategies for delayed intentions. In particular, we aimed to examine how optimal older adults are in a paradigm where they need to balance the

cost against the benefit of using reminders. Results supported previous laboratory studies (Henry et al., 2004; Uttl, 2008) that have shown an age-related decline in PM task performance. Consequently, it was optimal in our task for older adults to use more reminders than younger. We found that older participants did indeed set numerically more reminders than the younger group. However, whereas younger adults had a pro-reminder bias relative to the optimal strategy, replicating previous results (Ball et al., 2021; Engeler & Gilbert, 2020; Gilbert et al., 2020; Kirk et al., 2021), this bias was significantly reduced in the older group, whose reminder-setting behaviour did not differ significantly from the optimal strategy. Therefore, we found opposite results depending on whether one considers A) the absolute number of reminders (increased in older adults), or B) the pro-reminder bias relative to the optimal strategy (decreased in older adults). This shows that even in a situation where older adults make greater use of environmental cognitive support, they may nevertheless show a reduced preference for such support in comparison with younger adults. Consistent with this finding, Henry and colleagues (2012) and Scarampi and Gilbert (2021) found that older adults do not necessarily compensate for impaired memory performance when a reminder-setting strategy is available.

A second aim of this study was to explore younger and older adults' metacognitive judgements about their performance. Younger participants were particularly underconfident about their ability to perform the task with internal memory, and only slightly underconfident about their accuracy with external reminders. This means that they overestimated the benefit of reminders: the predicted difference between accuracy with versus without reminders was greater than the actual difference. This pattern was reversed in older adults, who were particularly underconfident in their ability with reminders but only slightly

underconfident in their ability with internal memory. Therefore, older adults underestimated the benefit of external reminders. This contrasting pattern between the two groups, with younger adults overestimating and older adults underestimating the benefit of reminders, could potentially account for the shift in behavioural strategies, with only the younger participants showing a pro-reminder bias. Consistent with this metacognitive account of offloading strategies, we found that individual differences, younger adults' metacognitive under- or over-confidence, were significantly correlated with their pro- or anti-reminder bias, replicating an effect seen in previous studies (Ball et al., 2021; Gilbert et al., 2020; Kirk et al., 2021). This correlation was nonsignificant in the older group. Potentially, this could reflect reduced metacognitive control in this group, i.e., the ability to translate metacognitive judgements into strategic reminder-setting decisions. However, the nonsignificant trend in the older group was in the same direction as the younger group, and the two correlations were not significantly different from each other, so it is not possible to draw strong conclusions from this.

While younger adults were highly underconfident about their unaided memory ability, older adults were better calibrated. This upward age-related shift in confidence, relative to actual performance, is consistent with other studies demonstrating increased overconfidence in older participants (Bruce et al., 1982; Cauvin et al., 2019; Connor et al., 1997; Scarampi & Gilbert, 2021; Soderstrom et al., 2012). The upward shift could potentially be explained by a failure to update metacognitive beliefs in line with an age-related decline in cognitive ability (cf. Knight et al., 2005 for a related phenomenon in the context of brain injury; Souchay, 2007 in the context of dementia). This could explain the failure of older adults to fully compensate for impaired memory for intentions when a reminder-setting strategy is

permitted (Scarampi & Gilbert, 2021). Consistent with this, in a study of PM in real life, Herzog et al. (2019) found that older people usually expected themselves to remember important PM tasks, relying highly on unaided memory, and did not view external strategies as a compensatory aid for memory decline. The practical implication of these findings is that cognitive offloading strategies could be optimised by designing metacognitive interventions that improve individuals' awareness of their true level of cognitive ability. This could particularly apply to older adults, who may otherwise fail to accumulate sufficient feedback to regulate their metacognitive beliefs (Touren & Hertzog, 2014). It could also be particularly relevant to the domain of delayed intentions, where individuals decide whether or not to offload intentions some time before the intended behaviour. This can lead to a long lag between the time at which a strategy is implemented and the time at which an individual can detect whether or not the strategy was effective. As a result of this long-time lag, feedback from task performance may fail to reliably update the metacognitive knowledge used for strategy selection.

Although older adults underestimated the benefit of reminders in this study, their offloading decisions were congruent with the optimal strategy. This pattern of results may appear somewhat surprising. If individuals selected offloading strategies based only on their metacognitive judgements, which underestimated the benefit of reminders, an anti-reminder bias would be predicted. Seeing as no such anti-reminder bias was observed, this suggests that one or more additional factors, other than metacognitive judgements, contributed to strategic offloading decisions. One potential factor that has been highlighted (Gilbert et al., 2022) is that individuals may prefer external reminders as a way of avoiding the cognitive effort associated with internal memory. Consistent with this, Sachdeva and Gilbert (2020) found that the

pro-reminder bias was reduced when participants had a financial incentive to behave optimally, which was hypothesised to increase participants' willingness to expend cognitive effort on the task.

While the present results are consistent with the view that older adults also choose to use external reminders as a means of avoiding cognitive effort, this tendency might be reduced if the act of setting reminders is itself seen as effortful. This possibility is congruent with previous research (Hertzog et al., 2012; Lineweaver et al., 2018) showing that older people choose strategies according to both their perceived difficulty and effectiveness, while younger people tend to consider effectiveness predominantly. Similarly, Hertzog et al. (2017) suggested that older people tended to use rote repetition rather than switching to more effective strategies due to the cognitive effort involved in switching strategies. Although intention offloading decreases cognitive effort as a result of removing the requirement to maintain an internal representation of intended behaviour, older people may consider the additional reminder-setting behaviour to be more effortful, for example due to the requirement to switch away from the ongoing task to physically set a reminder. Future research could investigate this factor by explicitly manipulating the effort associated with reminder-setting.

3.4.1 Limitations of this study

We note three main limitations of this study that could be addressed in future research. First, it is unclear to what extent the differences we observed between younger and older participants could reflect motivational effects, which could derive from differing reasons for taking part in the research to begin with (Ryan & Campbell, 2021). While we equalised the incentive structure of the task by using an explicit

point system which determined participants' payment at the end of the experiment, the effect of this could have differed between younger and older participants.

Previous research has shown that financial reward can improve younger adults' performance significantly (Aberle et al., 2010; Honeywell et al., 1997; Shah et al., 1998; Shum, 2004), however the effect on older people is less clear (Birkhill & Schaie, 1975; Strayer & Kramer, 1994; Touron et al., 2007). Therefore, it would be useful to investigate how far the effects reported here generalise across different tasks with a variety of incentive structures.

Second, our older participants were relatively well-educated (mean: 16 years), so the results may not generalise to the older population as a whole. The older participants in this study may potentially have chosen unaided memory because they were confident of their memory abilities or believed that staying mentally active helped to prevent cognitive decline. It would be helpful to investigate whether similar results hold across different levels of education and/or beliefs about the benefits of mental activity.

Finally, it is unclear how far age differences in cognitive offloading strategies are A) domain general, B) relate to specific aspects of cognition such as PM (e.g., due to age-differences in the prospective and/or retrospective components of PM), or C) are even more fine grained than that. For example, results could potentially differ depending on whether reminder-setting is performed via a digital device or a more traditional approach such as written notes. Several studies propose that older people tend to avoid modern technology (Chen & Chan, 2014; Ostrom et al., 2013).

Therefore, the impact of attitudes towards technology, and the specific mechanism of intention offloading, could potentially play an important role in the age differences

reported here, however this is ultimately an empirical question.

3.5 Conclusion

We report four main findings in the present study. First, older adults were significantly less likely than younger adults to remember delayed intentions when they used internal memory. Second, when they were given the option to set external reminders, older adults did so more often. Third, the age-related decline in memory performance was substantially reduced, but not eliminated, when reminders were used (see also Scarampi & Gilbert, 2021 for a similar result). The main novelty of this study was that we used a method that distinguishes whether increased reminder-setting in older adults reflects A) an adaptive response to impaired unaided ability, versus B) a change in individuals' preference towards using internal memory versus external cognitive tools. Despite older adults' increased use of reminders overall, their bias towards reminders relative to the optimal strategy was reduced. Therefore, the increased use of reminders in older adults was attributable to their greater need for external memory support, rather than an increased pro-reminder bias. In contrast to the view that ageing is characterised by an increased preference for environmental cognitive support (Lindenberger & Mayr, 2013), our results show that in some situations older adults have a *reduced* preference for external cognitive support when one takes into account their level of performance when such support is not available. These results may be attributable to age-related change in metacognitive evaluations. Therefore, we suggest that metacognitive interventions could be an effective means for optimising individuals' use of cognitive tools, across the lifespan.

Experiment 3: Effects of Delay Length and Metacognition on Strategic Reminder Setting in Time-based Prospective Memory

4.1 Introduction

Experiment 1 and 2 investigated whether metacognitive advice and aging processes are key factors in offloading intentions in event-based PM tasks. In Experiment 3, we evaluated two possible factors which influence intention offloading in a time-based PM task. To perform a time-based task, people have to fulfill an intention at a designated time or after a particular duration. For example, if you decide to make a cake, and put it into an oven, you need to remember when to take it out. Given that there is no event associated with the intended action, the retrieval processes of time-based intention might be different from event-based tasks, and depend more on implicit cues and self-initiated thoughts (Einstein & McDaniel, 1990; Kvavilashvili & Fisher, 2007).

In daily life, people may rely on time-monitoring to determine when to fulfil an intended action, or use external aids such as setting an alarm to complete tasks successfully. Using reminders means people would not be required to monitor time while waiting for a specific time, and focus more on the ongoing work (e.g. doing housework while waiting the cake to bake). In other words, setting reminders could free up the cognitive resources until the time people need to retrieve and fulfil intentions, which helps people to execute both ongoing and PM tasks more efficiently. Explicit reminders or cues could not only improve the performance on the

time-based PM tasks, but also the performance on the ongoing tasks (Cook et al., 2007). Despite these benefits, people do not always set reminders for everyday time-based PM tasks. Factors impacting whether people set reminders in a time-based PM task have seldom been investigated. In the current study, we investigated two time-related factors: delay length and time-monitoring behaviour. It is straightforward to consider the effect of delay length on intention offloading in a time-based PM task. For example, people might set an alarm when food needs to be taken out of an oven after one hour, but not when it takes only five minutes. Next, we evaluated intention offloading with attention to differences in time-monitoring behaviour associated with the visibility of the clock. Harris and Wilkins (1982) proposed the test-wait-test-exit model suggesting people would repeatedly check whether it is the appropriate time to fulfill the intention. If it is not, they will wait until checking (i.e. time-monitoring) again. In the end, people perform the delayed intention when the designated time is reached.

Before considering whether the two factors would influence intention offloading in a time-based PM task, a related question would emerge: would the extent to which people fulfill delay intentions be influenced by delay length and time monitoring? First, while abundant evidence from the research of retrospective memory (Brown, 1958; Peterson & Peterson, 1959) suggests memory performance declines due to a longer delay, very few studies have evaluated PM decay over time. In a study of Einstein et al. (2003), participants were instructed to execute ongoing tasks, and also required to press a specific key when seeing a red screen, but they should do it after the current task was finished. The tasks lasted for different periods (i.e. 5s, 15s and 40s). There was no significant performance difference across the three delay lengths, which suggested no significant PM decline over the short

interval (i.e. 40s). However, seeing as the shift in different tasks served as a cue for the intended action, the paradigm should be defined as an event-based PM task. Moreover, given the PM responses were scored as correct ones whenever participant pressed the key within the one-minute execution period, the PM accuracy was defined loosely. Conte and McBride (2018) required participants to perform a PM task with different delays (i.e. 1, 3 or 6 mins). Half the participants completed an event-based PM task, and the other half completed a time-based PM task. While the longer delays decreased the performance on the event-based PM task, no significant effect of the delay length on the performance was found on the time-based task. Further studies are needed to clarify whether PM declines over time on a time-based PM task.

Second, considerable time-based PM literature (Einstein et al., 1995; Harris & Wilkins, 1982; Henry et al., 2004; Mäntylä et al., 2009; Maylor, 1990) has suggested that time-monitoring behaviour could be an important predictor to PM performance. In order to evaluate the effect of time-monitoring on PM performance, an experiment manipulation to display persistently or hide the clock could be used. A clock displayed persistently acts as an external cue to induce the retrieval of PM target, which reduces cognitive load of PM tasks. In addition, people can check the time without spending additional cognitive or physical effort. In contrast, when the clock is out of sight, participant need to sustain a memory representation of monitoring time, and self-initiated the clock-checks (Harris & Wilkins, 1982; Mioni et al., 2020; Mioni & Stablum, 2014). Moreover, people have to inhibit an ongoing activity and switch their attention to the time-monitoring behaviour (Harris & Wilkins, 1982; Mioni et al., 2020; Mioni & Stablum, 2014). Given the PM task with a hidden clock is more demanding for cognitive effort, prior studies have revealed the

PM performance is worse than when the clock is constantly displayed (Aberle et al., 2010; Mioni et al., 2020).

On the other hand, delay length and time monitoring could influence the decision about intention offloading regardless of the extent to which the factors contribute to the performance. One possibility is the use of reminders is predicted by the metacognitive judgment in PM abilities. Previous PM research (Gilbert, 2015a, 2015b; Gilbert et al., 2020) and prior two experiments (i.e. Experiment 1&2) have demonstrated people's metacognitive judgment of the memory performance contributes to the decision about setting reminders in event-based tasks. If the metacognitive judgment varies with the lengths of the delays or the time monitoring behaviour, the view anticipates people change the reminder usage accordingly. For example, people might be more inclined to set reminders in the PM task with a longer delay seeing as they believe the task with a longer delay is more difficult

In order to investigate the issues outlined above, we designed two experiments. In the first experiment (i.e. Experiment 3a), participants fulfilled the time-based PM intentions after different lengths of short delays (10s, 20s, and 30s). In half of the trials, participant performed the unaided task while in the other half, they were permitted to set reminders. Before and after the experiment, participants provided a subjective rating of how well they expected they would perform or they thought they performed. The metacognitive judgements were given separately for different delay lengths. This allowed participants' reminder setting to be related to objective performance measures and metacognitive judgements separately for different delay lengths.

The aim of the second experiment (i.e. Experiment 3b), was to evaluate the

relationship between reminder usage and time-monitoring behaviour. Time-monitoring was manipulated by utilising clock revealability as a between-subject factor. Participants were divided into two groups, hidden-clock and persistent-on. The hidden-clock group had to press a specific button to check the time while a clock was in the constant view of participants in the persistent-on group.

In the current study, we developed a time-based PM task allowing us to evaluate the effect of delay length and time-monitoring behaviour on reminder usage. We would assess these questions: 1) Do individuals offload intentions more often when the delay of fulfilling intentions is longer? 2) Do individuals offload intentions more often when time-monitoring is more cognitively demanding? 3) If the effects on intention offloading do be shown, to what extent can they be attributed to metacognitive factors? Before data collection, we pre-registered our hypotheses, experimental procedure and analysis plan for Experiment 3a (<https://osf.io/snf8p/>), and Experiment 3b (<https://osf.io/jp462/>).

4.2 Experiment 3a

4.2.1 Method

The design of this experiment was a 3×2 factorial in which the delay lengths (i.e. 10 s, 20 s, 30 s) and the offload conditions (unaided and reminder) were varied within subjects. In all conditions, participants performed the on-going task, 2-back version of the N-back working memory paradigm. In this task, stimuli were letters of the alphabet. One letter at a time was presented at the middle of the screen. Participants responded to each stimulus by pressing the X key for targets (20 % of trials) which were letters identical to the one presented 2 trials back, or by pressing

the Z key for non-targets.

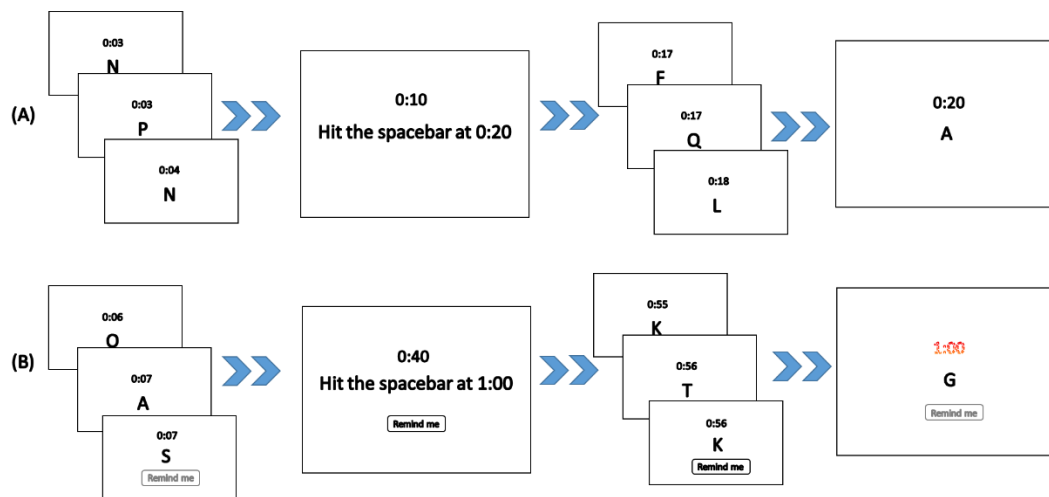
A digital clock was displayed above the letters, and started running when participants began the 2-back task (see Figure 4.1). Participants were told that during the ongoing task, they would be shown the instruction of the PM task. For example, a message, “Hit the spacebar at 0: 40”, was presented on the screen. The clock showed the current time was 0:30 so participants should keep going with the 2-back task. When the clock reached the designated time, 0:40, they then pressed the spacebar. In other words, the length of the delay was 10 seconds in the PM task. The response was counted as correct when participants pressed the spacebar during the window of two seconds (e.g. 0:38 to 0:42). In order to reduce the chance of ceiling performance on the PM task, the instructions before the experimental trials emphasised the importance of the 2-back task. We also encouraged participants to attend to the 2-back task by giving them an additional bonus. Participants were told they got one point for each correct response in the 2-back task. If their scores were in the top half, they would receive an extra £1.

There were 6 blocks in the experiment, three blocks of the unaided condition and three blocks of the reminder condition. In the unaided condition, participants depended on their own memory to perform the PM task while in the reminder condition, they could set reminders by clicking a button which shows “remind me”. This means the clock would flash to remind them during the four-second PM response window. It was completely up to the participants whether to use the button or not. The blocks of the two conditions (unaided and reminder) alternated, and the order of the two conditions counterbalanced across the participants. Each block included 12 time-based PM intentions in total. Following each designated PM

response time, there was a 10 s delay until the next instruction was presented. Participants completed four PM intentions for each delay length (i.e. 10 s, 20 s, 30 s), and the order of the three delay lengths was determined randomly. Each block lasted a total of 365 seconds.

Figure 4.1

A schematic illustration of the PM task with a 20 s delay in the unaided and reminder conditions.



Note. A digital clock above the letters starts running when the participants begin a block. The participants are instructed to perform the 2-back task, and then the screen shows the instruction of a PM task.

(A) The unaided condition: The participants should keep performing the 2-back task. When the clock reaches the designated time, 0:20, they should remember to push the spacebar.

(B) The reminder condition: After the instruction of the PM task is shown, the participants could choose to click the button which shows “remind me”. They should keep going with the 2-back task. The clock will flash to remind the participants when it is nearly the designated time.

4.2.2 Procedure

Participants first had a short practice session of the 2-back task and the PM task.

For each delay length (i.e. 10 s, 20 s, 30 s), there was 3 practice trials of the PM task. They would be required to repeat the block if there was no correct PM response. After practice trials, they were asked to predict what percentage of PM intentions (from zero to 100) they would fulfill successfully. The predictions were made separately for each delay length (i.e. 10 s, 20 s, 30 s). Then they were told there would sometimes be a reminder button, and practiced how to use it once. The trial would be repeated if they did not set the reminder. Before beginning the experiment, participants were instructed that they were free to choose using the button or not. After the main experiment, participants were asked to report what percentage of PM tasks they thought they successfully performed, using unaided memory. The post-task confidence ratings were also made separately for each delay length. Participants received £7 payment after completing the experiment, and an extra bonus of £1 if their scores of the ongoing task were in the top half.

A demonstration of the full experiment including all instructions and practice can be viewed here: <http://cognitiveoffloading.net/PT1>

4.2.3 Participants

Participants were recruited via the Prolific website, and took part by accessing a web-link that was provided to them. The study was approved by the UCL Research Ethics Committee (1584/002), and informed consent was obtained from all participants.

A priori power analysis was conducted based on a previous event-based PM study corresponding to the main investigation of the current study, which is evaluating the relationship between the individual difference of the metacognitive

evaluation and the individual difference in the reminder usage. (Kirk et al., 2020) examined whether participants could set reminders optimally while weighing the benefits/costs of setting reminders. The significant correlation coefficient between the metacognitive bias (i.e. the difference between confidence and actual accuracy) and the reminder bias was -0.34. This was conceptually related to our main purpose of finding the effect of the metacognitive evaluation on the reminder usage. In order to achieve 80% power to replicate the effect in the study (one-tailed, $\alpha = .05$), a sample size of 52 was required (G*Power 3.1). If participants were excluded due to the exclusion criteria below, they were replaced until a final sample size of 52 is reached.

Exclusion Criteria:

1. Accuracy below 70% on non-match trials of the 2-back task.
2. Accuracy below 40% on match trials of the 2-back task.
3. Mean PM accuracy below 10%.
4. Mean PM accuracy more than 2.5 absolute deviations from the median of all participants (Leys et al., 2013).
5. Mean PM confidence more than 2.5 absolute deviations from the median of all participants (Leys et al., 2013).

Eleven participants were excluded due to our pre-registered criteria (<https://osf.io/snf8p/>). Nine were removed due to accuracy below 70% on non-match trials of the 2-back task. Two participants were excluded because their mean PM confidence was more than 2.5 standard deviations from the median of all participants. No participant met any of the other exclusion criteria.

4.2.4 Dependent Measures

- A. These measures were calculated separately for each offload condition (unaided/reminder) as well as each delay length (10 s/20 s/30 s).
- i ‧ PM accuracy (PA): The mean proportion of PM trials where the spacebar was pressed within the 2-second instructed response window.
 - ii ‧ PM accuracy without the repeat hits (PAW): Mean proportion of PM trials where the spacebar was pressed within the four-second response window and was not pressed prior to this. Participants occasionally experienced inadvertent forgetfulness of the instructed time and resorted to repeatedly pressing the spacebar. To account for this, this measure excluded trials where the participant repeatedly pressed the spacebar.
 - iii ‧ False hits: The mean number of times participants pressed the spacebar outside the instructed time.
 - iv ‧ Missing hits: The proportion of trials where the spacebar was not pressed at all.
- B. The measures were calculated separately for each delay length (10 s/20 s/30 s).
- i ‧ Reminder usage: The proportion of trials when participants set reminders by pressing the reminder button in the reminder condition.
 - ii ‧ Confidence: The average of the pre-task and post-task confidence ratings about the PM accuracy in the unaided condition.
 - iii ‧ Metacognitive bias: The difference between the subjective confidence and the actual accuracy (i.e. PAW).

$$\text{Metacognitive bias} = \text{Confidence} - \text{PAW}$$

- C. The measures on the ongoing task were calculated separately for each offload

condition (unaided/reminder), as well as each PM target condition (no-target/ 10 s/ 20 s/ 30 s). In the no-target condition, participants did not maintain a PM intention because the ongoing tasks were performed before the designated time of the next PM task was presented. Before analysing data from the ongoing task, we excluded the following trials: a) the first trial of each block, b) the trial before and after any spacebar press, c) the trial after any PM instruction, and d) any trial with RT below 150ms or above 3000ms.

- i. Accuracy of the match trials on the 2-back task(A_m): The mean target accuracy (i.e. proportion of the correct responses to the target letters) on the 2-back task
- ii. Accuracy of the non-match trials on the 2-back task(A_n): The mean non-target accuracy (i.e. proportion of the correct responses to the non-target letters) on the 2-back task
- iii. D-prime on the 2-back task(d') : The difference between the z-transforms of A_m and $(1-A_n)$:
$$d' = z(A_m) - z(1-A_n)$$
- iv. Reaction time of the match trials on the 2-back task(R_m): The mean reaction time of the match trials on the 2-back task
- v. Reaction time of the non-match trials on the 2-back task(R_n): The mean reaction time of the non-match trials on the 2-back task

4.2.5 Results

All analyses were conducted in accordance with the pre-registered plans, except where stated (<https://osf.io/snf8p/>). All data analyses were conducted using R version 4.1.2 (R Core, 2021). See Table 4.1 and Figure 4.2 for a summary of results.

First, the data of the PM performance (PA, PAW, false hits and missing hits) using unaided memory were submitted to 1×3 ANOVAs including a within-subjects variable, the delay length (10s, 20s, 30s). As shown in Table 4.1, the analysis revealed that the delay length had no significant effect on the PM accuracy and missing hits. However, in terms of the false hits (see Table 4.1), the longer the delay length was, the more frequently participants pressed the spacebar outside the instructed time. In line with the results of the false hits, PAW (i.e. excluding the accurate PM responses if there were false hits ahead of the responses) was reduced in the trials with a longer delay. One possible explanation is that participants who forget the instructed time periodically pressed the space bar (e.g. every 10 s), and therefore the longer delay duration meant there would be more false hits and less PAW. The results suggest when participants were allowed to make multiple PM responses, there was no significant evidence that PM performance declined over time. However, if only the first response was considered, there was clear evidence for a decline in PM accuracy with longer delays.

Second, we conducted a repeated-measures ANOVA on reminder usage in the three delay conditions. The analysis suggested a significant effect of the delay length (see Table 4.1 & Figure 4.2). Post-hoc analyses revealed that all the pairwise differences, between three delays, were significantly different (10s/20s: $t(51) = -4.65$, $p < .001$; 10s/30s: $t(51) = -6.22$, $p < .001$; 20s/30s: $t(51) = -3.56$, $p < .001$). The results showed that participants set reminders more often at longer delays.

This brings us to the next point, the metacognitive judgements about PM performance. A repeated-measures ANOVA conducted on confidence (see Table 4.1 & Figure 4.2) showed the confidence scores (averaged across the two judgements)

was significantly different between different delay lengths. Post-hoc pairwise paired t-tests revealed the differences in confidence between 30s and the other two delay lengths were significantly different, but not between 10s and 20s (10s/20s: $t(51) = 1.73, p = .090$; 10s/30s: $t(51) = 3.92, p < .001$; 20s/30s: $t(51) = 4.03, p < .001$). Moreover, the metacognitive bias (i.e. the difference between confidence and PAW) was also compared between the three delay lengths by using a repeated-measures ANOVA (see Table 4.1). The analysis and post-hoc pairwise paired t-tests showed the differences in metacognitive bias was significant between 10s and the other two delay lengths, but not between 20s and 30s (10s/20s: $t(51) = -3.5, p < .001$; 10s/30s: $t(51) = -3.14, p < .001$; 20s/30s: $t(51) = -0.94, p = .354$). Thus, participants were less confident of the PM performance at longer delays, but the metacognitive bias was reduced.

4.2.5.1 Metacognitive Judgment and Reminder Usage

To examine whether the individual differences in metacognitive judgement would influence the individual differences in reminder setting, we calculated each participant's confidence, reminder usage, unaided PM accuracy, and metacognitive bias by averaging each measure across three delay lengths. The results of one-tailed Pearson correlations showed the reminder usage was not correlated significantly with confidence, metacognitive bias and PM performance (confidence: $r(50) = -.11, p = .216$; metacognitive bias: $r(50) = -.08, p = .291$; PM accuracy: $r(50) = -.02, p = .450$; PAW: $r(50) = -.03, p = .404$). Therefore, no significant evidence was found supporting that participants who were less confident about performance used more reminders overall. Neither did participants who had worse PM performance use more reminders.

Table 4.1.

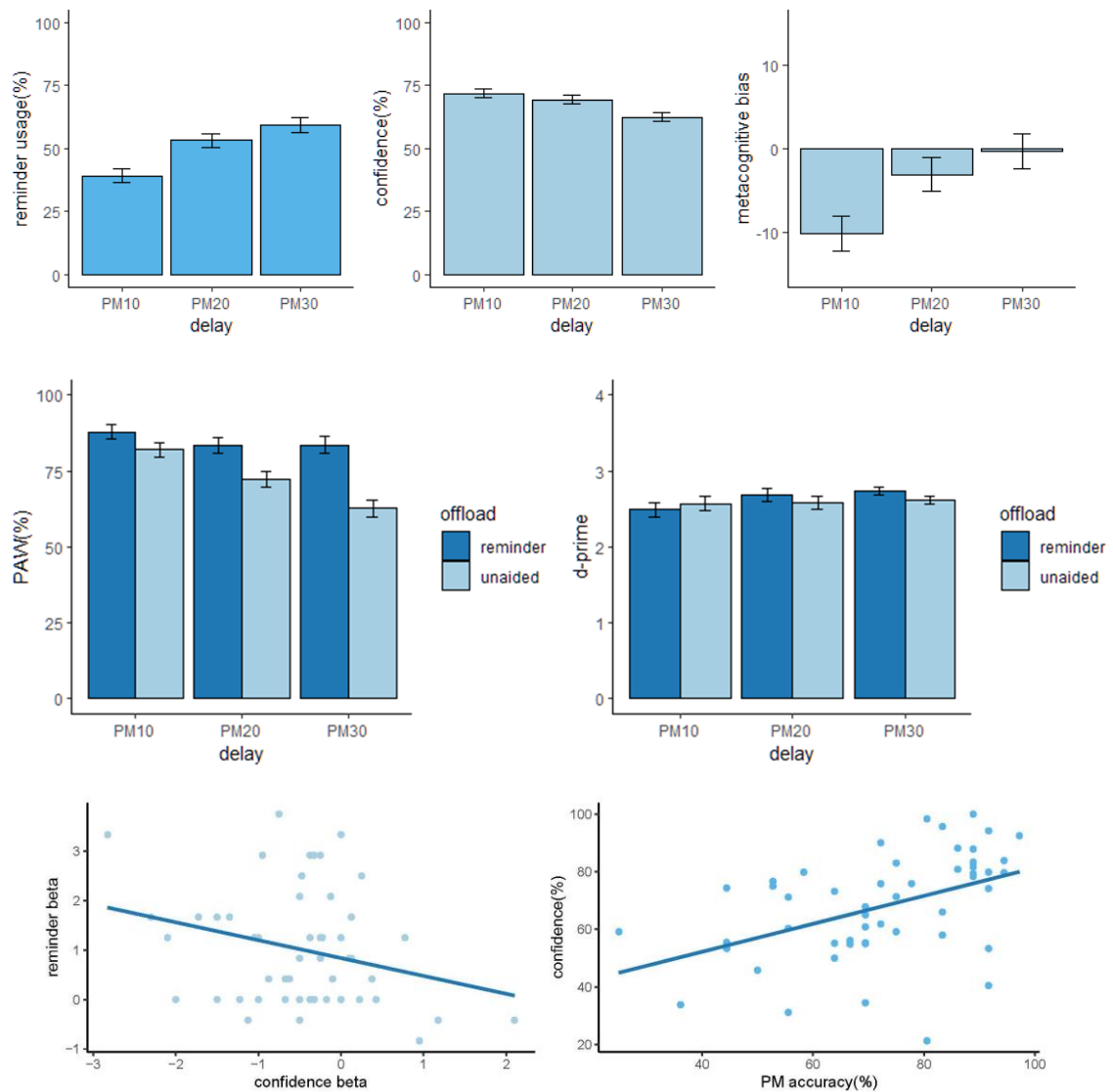
Average scores of the PM measures for each delay length and analyses of variance examining the effects of delay length on the measure scores (F-values, significance levels and partial eta-square). Standard errors of mean are shown in parentheses.

Measure	Delay			F	p	η^2_p
	10 s	20 s	30 s			
PA (%)	86.7 (13.7)	86.4 (14.3)	85.3 (17.1)	.44	.646	.01
PAW (%)	82.1 (16.2)	72.4 (17.8)	62.8 (23.2)	36.87	< .001***	.42
False hits	1.48 (1.96)	3.73 (3.85)	6.81 (6.29)	32.41	< .001***	.39
Missing hits (%)	1.04 (1.41)	.89 (1.22)	.75 (1.2)	2.2	.116	.04
Confidence (%)	72 (21.4)	69.4 (18.2)	62.5 (19.7)	13.08	< .001***	.20
Metacognitive bias (%)	-10.1 (20.3)	-3.05 (20.9)	-.34 (23.3)	6.87	.003**	.12
Reminder usage (%)	39.3 (38.5)	53.2 (39.1)	59.5 (38.6)	28.34	< .001***	.36

Note: The calculation of the measures except reminder usage only includes the PM trials in the unaided condition.

Figure 4.2

Results from the PM task and ongoing task in Experiment 3a



Note.

Upper row: Error bars represent within-subject confidence intervals. The error bars do not overlap with each other, which means $p < .05$.

Left: the reminder usage in the PM trials with 10s, 20s, and 30s delay.

Middle: the confidence in the PM trials with 10s, 20s, and 30s delay.

Right: the metacognitive bias in the PM trials with 10s, 20s, and 30s delay.

Middle row: Error bars represent within-subject confidence intervals (comparing the reminder and unaided conditions). The error bars do not overlap with each other, which means $p < .05$.

Left: The PM performance (i.e. PAW) as a function of Delay (10s, 20s, 30s) and Offload (unaided, reminder).

Right: the ongoing task performance (i.e. d-prime) in the PM trials as a function of Delay (10s, 20s, 30s) and Offload (unaided, reminder).

Lower row:

Left: Relationship between confidence beta (i.e. the relationship between the confidences at the three delays) and reminder beta (i.e. the relationship between the reminder usages at the three delays). Both confidence and reminder beta values were obtained by calculating a separate regression analysis for each participant. The blue line indicates linear regression for the relationship between the two variables across all individuals.

Right: Relationship between confidence (averaged across the three delays) and PM performance (averaged across the three delays). The blue line indicates linear regression for the relationship between the two variables across all individuals.

4.2.5.2 Modification across Three Delays

Despite no significant evidence showing participants who were generally less confident across the three delays used more reminders overall, we turned to evaluate the role of metacognitive judgment in modifying reminder usage at different delays. In the pre-registered analysis, we proposed to perform a separate regression analysis for each participant, conducting a linear regression on reminder usage with a regressor, the confidence at the three durations (averaged across the two judgements).

However, the analysis did not investigate whether the extent to which the delay length impacted metacognitive judgements predicted the effect of the delay length on reminder setting. Therefore, instead of doing this, we decided to conduct additional analyses which were not pre-registered. We first conducted a simple linear regression for each subject to evaluate how delay duration predicted the confidences (averaged across the two judgements), and another simple linear regression to evaluate how delay duration predicted reminder usage. A significant correlation ($r(50) = -.27, p = .027$) between the two resulting beta values was demonstrated by a

one-tailed Pearson correlation. Therefore, participants who had a steeper drop in confidence at longer delays tended to sharply increase reminder usage at longer delays.

We also evaluated the relationship between reminder usage and PM performance (i.e. PAW) across three delays. First, we conducted a simple linear regression for each subject to evaluate how delay duration predicted the PAW, and another simple linear regression to evaluate how delay duration predicted reminder usage. Then a one-tailed Pearson correlation between the two resulting beta values was conducted, which did not show significant negative relationship ($r(50) = .26$). Therefore, participants whose performance dropped more deeply at longer delays did not increase reminder usage more sharply at longer delays.

Next, we evaluated the relationship between confidence and PM performance (i.e. PAW) across three delays. First, we conducted a simple linear regression for each subject on PAW with a regressor, delay duration, and another simple linear regression on the confidences (averaged across the two judgements) with a regressor, delay duration. Then a one-tailed Pearson correlation between the two resulting beta values was conducted, which did not show significant negative relationship ($r(50) = .14$). Therefore, participants whose performance dropped more deeply at longer delays did not have more sharply decreased confidence at longer delays.

4.2.5.3 PM Performance Improvement

In the next analysis, we assessed the effect of offloading intentions by comparing PM performance on both the unaided and reminder trials (see Figure 4.2 Middle Row). Specifically, repeated-measures 2x3 ANOVAs on the measures of PM

performance (i.e. PA, PAW, false hits, missing hits) with factors Offload (unaided, reminder) and Delay (10s, 20s, and 30s) were performed. The analysis revealed that setting reminders significantly improved the PM accuracy ($F(1, 51) = 22.18, p < .001, \eta^2_p = .3$). Further analyses showed the effect of offloading intentions was significant at 20s ($p < .001$) and 30s delay ($p < .001$), but not at 10s delay ($p = .073$). There was also a significant interaction between Offload and Delay ($F(2,102) = 3.14, p = .048, \eta^2_p = .06$). As can be seen in Figure 4.2, the effect of offloading on the PM accuracy was more prominent at longer delay. Similar results were found in ANOVA analyses conducted on false hits (Offload: $F(1,51) = 26.33, p < .001, \eta^2_p = .34$; Delay \times Offload: $F(2,102) = 19.54, p < .001, \eta^2_p = .28$) and PAW (Offload: $F(1,51) = 37.11, p < .001, \eta^2_p = .42$; Delay \times Offload: $F(2,102) = 15.71, p < .001, \eta^2_p = .24$). The analyses conducted on the missing hits showed the main effect of Offload also reached significant ($F(1,51) = 7.53, p < .008, \eta^2_p = .13$), but the Delay \times Offload interaction was not significant ($F(2,102) = .11, p = .897, \eta^2_p < .01$). Thus, the results overall suggested offloading intentions improved PM performance. More importantly, PM performance was improved more profoundly at longer delay, which can be explained by the above finding that participants offloaded intentions more often at longer delay.

4.2.5.4 Ongoing Task Performance

Finally, we investigated whether the performance of the ongoing task differed among the three delay lengths, and was affected by reminder usage (see Figure 4.2 Middle Row). A repeated-measures ANOVA analysing the d' measure from the 2-back task with factors, Offload (unaided, reminder) and Delay (no-target, 10s, 20s, 30s), revealed the main effects of Delay were significant ($F(3,153) = 3.83, p = .001, \eta^2_p = .07$). Pairwise comparison, using paired t-test showed that the mean d' (no-target: $M = 2.65, SD = .76$; 10s: $M = 2.53, SD = .77$; 20s: $M = 2.63, SD = .79$; 30s: M

=2.68, $SD = .77$) was significantly different between 10s delay and 30s delay ($p = .004$). Neither the Offload main effects nor the Offload \times Delay interaction was significant. We also analysed the RT measure with factors, Offload (unaided, reminder), Match (match, nonmatch), and Delay (no-target, 10s, 20s, 30s). There were significant main effects of Delay ($F(3, 153) = 15.33, p < .001, \eta^2_p = .23$) and Match ($F(1, 51) = 33.66, p < .001, \eta^2_p = .40$), but the main effect of Offload was not significant ($F(1, 51) = .39, p = .54, \eta^2_p < .01$). The three-way interaction was not significant ($F(2.5, 127.28) = 1.16, p = .323, \eta^2_p = .02$), nor were the Offload \times Match interaction ($F(1, 51) = .73, p = .398, \eta^2_p = .01$), or the Offload \times Delay interaction ($F(2.36, 120.48) = .76, p = .49, \eta^2_p = .02$). However, the Delay \times Match interaction was significant ($F(2.03, 103.46) = 3.11, p = .048, \eta^2_p = .06$). Response times were significantly shorter on the match trials compared with nonmatch trials (match: mean=645ms, $SD = 150$; nonmatch: mean = 695ms, $SD = 141$). On the match trials, responses times were significantly shorter when participants were performing ongoing task only (no-target: mean = 624, $SD = 143$; 10s: mean= 646, $SD = 149$; 20s: mean= 659, $SD = 157$; 30s: mean= 651, $SD = 151$). These results suggest that the d' -prime and response time of the 2-back task were not significantly improved by offloading. In terms of delay length, the performance (i.e. d') of the ongoing task dropped when participants were performing at shorter delays.

4.2.5.5 Additional Non-preregistered Analyses

Some additional statistical analyses the pre-registration did not include were performed. First, a one-tailed one-sample t-test was conducted to assess whether metacognitive bias (averaged across the delay lengths) was less than zero. The results suggested participants were significantly underconfident of the unaided memory performance ($t(51) = -1.77, p = .042, d = -.24$).

Second, one-tailed Pearson correlations showed confidence (averaged across delays) was correlated significantly with unaided PM performance (averaged across three delays) (PM accuracy: $r(50) = .40, p < .001$; PAW: $r(50) = .45, p < .001$, see Figure 4.2 Lower Row). Also, confidence was negatively correlated with false hits ($r(50) = -.46, p < .001$). Accordingly, participants generally had fair metacognitive judgement about the PM performance including both the correct and false hits.

Third, some participants might stick to one strategy such as using unaided memory rather than choosing strategies adaptively. The presence of these participants in the sample might mask a relationship between confidence and reminder setting amongst those participants who did sometimes set reminders. After excluding the participants seldom setting reminders (i.e. equal or less than once), a one-tailed Pearson correlation conducted on the remainder (43/52 participants) showed a significant negative correlation between confidence and reminder usage ($r(41) = -.35, p = .010$).

Finally, we used lmer (Kuznetsova et al., 2017) to perform a linear mixed model analysis of the relationship between metacognitive judgement and reminder usage across three delays. As fixed effects, there were variables: a) delay length (i.e. 10 s, 20 s, and 30 s) b) confidence (averaged across the two judgements) c) PM performance (i.e. PAW) in the model. As a random effect, there were intercepts for participants. Likelihood ratio tests of the full model with the effect in question against the model without the effect in question were conducted (confidence: $\chi^2(1) = 7.05, p = .007$; PM performance: $\chi^2(1) = 1.43, p = .23$). The results suggested confidence predicted unique variance of reminder usage ($\beta = -.30, SE = .12$).

4.3 Experiment 3b

4.3.1 Method and Procedure

The design of this experiment was a $3 \times 2 \times 2$ factorial in which the delay lengths (i.e. 10 s, 20 s, 30 s) and the reminder conditions (unaided and reminder) were varied within subjects, and the clock revealability was varied between subjects.

The overall experiment was identical to Experiment 3a except few changes. First, participants were told to press the spacebar only once for each PM task and they were told that only the first spacebar response for each PM task was counted, which discouraged participants from pressing the spacebar repetitively when they forgot the instructed time. Second, participants were randomly assigned to one of the two groups, persistent-clock and hidden-clock. The groups differed as to whether the clock was persistently displayed or hidden. For the persistent-clock group, a digital clock was displayed above the letters, which was the same as Experiment 3a. For the hidden-clock group, no clock was displayed when participants began each block. To monitor for the passage of time, participants were told to press the M key at any time during the PM task. When they pressed the M key, the clock appeared above the letters for a duration of 2 seconds. In addition to the volitional time monitoring, the clock was also displayed when participants were instructed about the PM target time so they knew how long they should wait. The clock then disappeared after participants pressed the spacebar to keep going with the ongoing task. The last modified feature between Experiment 3a and Experiment 3b was that participants set reminders by clicking a button five times instead of once. The increased number of clicks helped to prevent a ceiling effect in reminder use. This also means that setting a reminder incurred a greater physical cost than checking the

hidden clock. Without this, there would be no logical reason for participants in the hidden-clock condition to ever use the clock-check mechanism rather than simply relying on a reminder. The button showed “remind me” and a number representing the remaining times participants should click.

In all other ways, the procedure matched that of Experiment 3a. A demonstration of the full experiment including all instructions and practice can be viewed here: <http://cognitiveoffloading.net/PT2>

4.3.2 Participants

Participants were recruited via the Prolific website, and took part by accessing a web-link that was provided to them. The study was approved by the UCL Research Ethics Committee (1584/002), and informed consent was obtained from all participants. The experiment took approximately one hour, and participants received a base payment, £7. In order to reduce the chance of ceiling performance on the PM task, we encouraged participants to attend to the 2-back task by giving them an additional bonus. Participants were told they got one point for each correct response in the 2-back task. If their scores were in the top half, they received an extra £1.

In Experiment 3a, 43 out of 52 participants used reminders not less than once across the PM tasks. For these participants, the significant correlation coefficient between confidence and reminder usage was $-.35$. In order to achieve 80% power to replicate the effect (one-tailed, $\alpha=.05$), a sample size of 48 would be required (G*Power 3.1) for each group (i.e. persistent-clock and hidden-clock). The total sample size for this experiment was therefore 96 participants with 48 participants in each group. If participants were excluded due to the exclusion criteria below, they

were replaced until a final sample size of 96 was reached.

Exclusion Criteria:

1. Accuracy below 70% on non-match trials of the 2-back task.
2. Accuracy below 40% on match trials of the 2-back task.
3. Mean PM accuracy below 10%.
4. Mean PM accuracy more than 2.5 median absolute deviations from the median of all participants in each group (Leys et al., 2013).
5. Mean PM confidence more than 2.5 median absolute deviations from the median of all participants in each group (Leys et al., 2013).
6. Reminder usage across all PM tasks equal to zero.

A total of 139 volunteers (68 persistent-on, 71 hidden-clock) to reach the sample of 96 participants (48 participants in each group) were tested. 43 participants were excluded due to our pre-registered exclusion criteria (<https://osf.io/snf8p/>). Thirteen (1 persistent-on, 12 hidden-clock) were removed due to accuracy below 70% on non-match trials of the 2-back task. Eight (3 persistent-on, 5 hidden-clock) were removed due to accuracy below 40% on match trials of the 2-back task. Eighteen (13 persistent-on, 5 hidden-clock) were removed due to mean PM accuracy below 10%. One participants (1 hidden-clock) were excluded because their mean PM accuracy was more than 2.5 standard deviations from the median of all participants. Three participants (3 persistent-on) were excluded because their mean PM confidence was more than 2.5 standard deviations from the median of all participants. No participant met any of the other exclusion criteria.

4.3.3 Dependent Measures

The dependent measures listed below were different from Experiment 3a, and everything else was the same.

- A. The measures were calculated separately for each offload condition (unaided/reminder) as well as each delay length (10 s/20 s/30 s).
 - i. PM accuracy (PA): The mean proportion of PM trials where the spacebar was pressed within the 4-second instructed response window. In Experiment 3a, some participants pressed the spacebar repeatedly for a PM task when they forgot an instructed time, which influenced the calculation of PM accuracy. Therefore, participants were told to press the spacebar only once for each PM target, and only the first response was counted. Therefore, there was only one measure here (i.e. no PAW).
- B. The measures are calculated for the clock-hidden group
 - i. Time-monitoring frequency: The number of clock-checking every 5-second period across the delay duration (i.e. 10s, 20s, and 30s). For example, in a trial with a 10s delay, time-monitoring frequency included two measurements: the number of clock checks was counted a) from 0s to 5s, and b) from 5s to 10s.
- C. The measures are calculated for the ongoing task: Except the exclusive criteria mentioned in Experiment 3a, we excluded the trial before and after M-key press.

4.3.4 Results

All analyses were conducted in accordance with the pre-registered plans, except

where stated (<https://osf.io/jp462/>). All data analyses were conducted using R version 4.1.2 (R Core Team, 2021). See Table 4.2, and Figure 4.3 for a summary of results.

4.3.4.1 PM Performance.

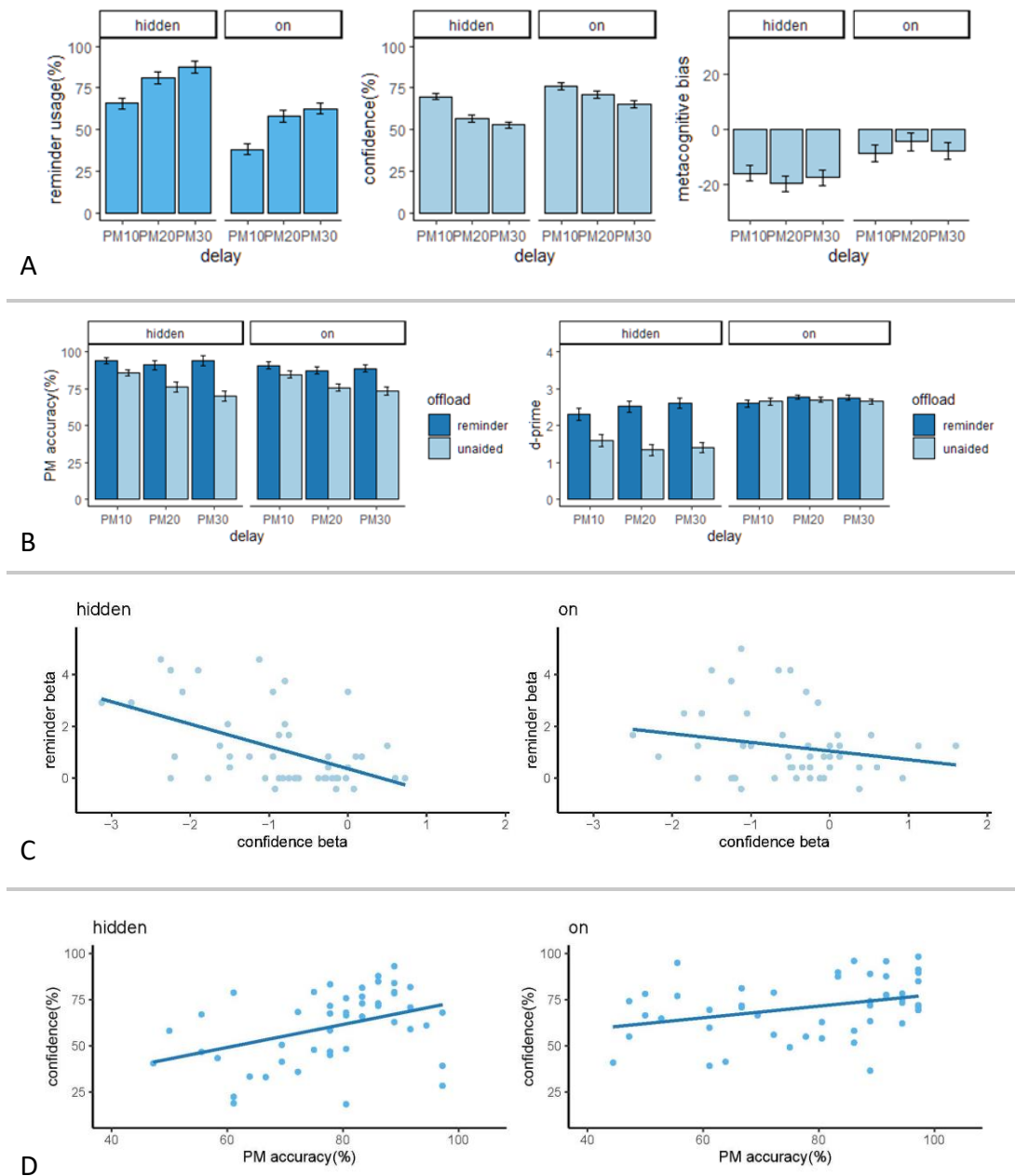
To examine the performance of PM, PM accuracy were subjected to a 3×2 mixed ANOVAs that included the between-subjects variable of Clock (persistent-on, hidden-clock) and the within-subjects variable of Delay (10s, 20s, 30s). The analysis revealed a main effect of Delay ($F(2, 188) = 25.01, p < .001, \eta^2_p = .21$). However, neither the effect of Clock nor the Delay \times Clock interaction (Clock : $F(1, 94) = .03, p = .855, \eta^2_p < .01$; Clock \times Delay: $F(2, 188) = .74, p = .478, \eta^2_p < .01$) were significant. The results indicates the PM performance declined over time, but no significant effect of the clock revealability was found.

4.3.4.2 Reminder Usage.

A 3×2 ANOVA on the reminder usage was conducted with Delay (10 s, 20 s, 30 s) as the within-subject factor and the Clock (persistent-on, hidden-clock) as the between-subject factor reveals that both the delay length and clock revealability had a significant main effect on reminder usage (Delay : $F(1.45, 136.24) = 51.93, p < .001, \eta^2_p = .36$; Clock: $F(1, 94) = 14.15, p < .001, \eta^2_p = .13$). There was no significant Delay \times Clock interaction ($F(1.45, 136.24) = .48, p = .556, \eta^2_p < .01$). The results suggested people set more reminders when the delay was longer or when the clock was not always visible.

Figure 4.3

Results from the PM task and ongoing task in Experiment 3b



Note.

(A) Error bars represent within-subject confidence intervals. The error bars do not overlap with each other, which means $p < .05$.

Left: the reminder usage in the PM trials as a function of Clock group (hidden-clock, persistent-on), and Delay (10s, 20s, 30s).

Middle: the confidence in the PM trials as a function of Clock group (hidden-clock, persistent-on), and Delay (10s, 20s, 30s).

Right: the metacognitive bias in the PM trials as a function of Clock group (hidden-

clock, persistent-on), and Delay (10s, 20s, 30s).

(B) Error bars represent within-subject confidence intervals (comparing the reminder and unaided conditions). The error bars do not overlap with each other, which means $p < .05$.

Left: The PM performance (i.e. PA) as a function of Clock group (hidden-clock, persistent-on), Delay (10s, 20s, 30s), and Offload (unaided, reminder).

Right: The ongoing task performance (i.e. d-prime) in the PM trials as a function of Clock group (hidden-clock, persistent-on), Delay (10s, 20s, 30s), and Offload (unaided, reminder).

(C) Relationship between confidence beta (i.e. the relationship between the confidences at the three delays) and reminder beta (i.e. the relationship between the reminder usages at the three delays). Both confidence and reminder beta values were obtained by calculating a separate regression analysis for each participant. The blue line indicates linear regression for the relationship between the two variables across all individuals.

Left: the hidden-clock group.

Right: the persistent-on group.

(D) Relationship between confidence (averaged across the three delays) and PM performance (averaged across the three delays). The blue line indicates linear regression for the relationship between the two variables across all individuals

Left: the hidden-clock group.

Right: the persistent-on group

4.3.4.3 Metacognitive Judgement.

We first conducted a 3×2 ANOVA on confidence (averaged across the two judgements) with Delay (10 s, 20 s, 30 s) as the within-subject factor and Clock (persistent-on, hidden-clock) as the between-subject factor. The analysis confirmed significant main effects of both Delay ($F(1.65, 155.34) = 42.73, p < .001, \eta^2_p = .31$) and Clock ($F(1, 94) = 9.18, p = .030, \eta^2_p = .09$). A significant interaction Delay \times Clock ($F(1.65, 155.34) = 4.25, p = .022, \eta^2_p = .04$) was also found. The results suggested that, participants were significantly less confident at longer delays in both groups.

However, the effect of delay on confidence was more prominent in the hidden-clock group. Then, a similar 3×2 ANOVA conducted on metacognitive bias (i.e. the

difference between confidence and PM accuracy) revealed a significant main effect of Clock ($F(1, 94) = 6.69, p = .007, \eta^2_p = .08$). However, the main effect of Delay ($F(2, 188) = .03, p = .974, \eta^2_p < .01$) and the interaction Delay \times Clock ($F(2, 188) = 1.47, p = .233, \eta^2_p = .02$) were not significant. Thus, the hidden-clock participants were more underconfident compared with the persistent-on participants. While the hidden-clock participants were increasingly underconfident at longer delays, the effect of delay was not significant in the persistent-on group.

4.3.4.4 Metacognitive Judgment and Reminder Usage

After examining metacognitive judgment at different delays, we evaluated the relationship between reminder usage and metacognitive judgement by conducting a linear regression on the reminder usage score (averaged across all delays), with regressors: A) mean confidence (averaged across all judgements, and mean-centred), B) mean PM accuracy (averaged across all delays), and C) clock revealability (coded as persistent-on = 1, hidden-clock = -1). This investigated unique variance attributable to each factor. The analyses of the beta coefficients for confidence and PM accuracy were one-tailed, based on the prediction that there should be a negative relationship between reminder use and each of the measures. Regression analyses showed that clock revealability predicted unique variance of reminder usage significantly, but confidence and PM accuracy did not (confidence: $\beta = .13, p = .742$; PM accuracy: $\beta = -.05, p = .838$; clock revealability: $\beta = -13.35, p < .001$).

Given no unique variance attributed to confidence and PM accuracy was found, we next performed a simple linear regression on reminder usage score (averaged across all delays) with regressors: A) mean confidence (averaged across all judgements, and mean-centred), B) clock revealability (coded as persistent-on = 1,

hidden-clock = -1), and C) confidence \times clock revealability interaction. The one-tailed analyses of beta coefficients showed only clock revealability was significant regressor (confidence: $\beta = .06, p = .617$; clock revealability: $\beta = -12.94, p < .001$ confidence \times clock revealability: $\beta = -.25, p = .096$). Additionally, for each group, one-tailed Pearson correlations revealed the correlation between confidence and reminder usage was not significant (persistent-on group: $r(46) = -.09, p = .283$; hidden-clock group: $r(46) = .22, p = .930$). The results revealed no significant evidence supporting individuals' average metacognitive judgment predicted the decision about the offloading strategy.

Regarding the effect of metacognitive bias (i.e. the difference between confidence and PM accuracy), we repeated the above regression analyses using metacognitive bias instead of confidence, and the corresponding analyses were also one-tailed. First, a regression analysis was conducted on reminder usage (averaged across all delays) with regressors: A) metacognitive bias, B) PM accuracy, and C) clock revealability (coded as persistent-on = 1, hidden-clock = -1). Only clock revealability significantly predicted reminder usage (metacognitive bias: $\beta = .134, p = .742$; PM accuracy: $\beta = .09, p = .634$; clock revealability: $\beta = -13.35, p < .001$). Next, a regression analysis was conducted on reminder usage (averaged across all delays) with regressors: A) metacognitive bias, B) clock revealability (coded as persistent-on = 1, hidden-clock = -1), and C) metacognitive bias \times clock revealability interaction. Only clock revealability significantly predict reminder usage (metacognitive bias: $\beta = .106, p = .723$; clock revealability: $\beta = -17.32, p < .001$; metacognitive bias \times clock revealability: $\beta = -.34, p = .064$). Finally, one-tailed Pearson correlations revealed for either group, the correlation between metacognitive bias and reminder usage was not significant (persistent-on group: $p = .209$; hidden-clock: $p = .975$).

The overall results indicated no significant relationship between averaged metacognitive measures (judgment and bias) of each participant (averaged across three delays) and averaged reminder usage (averaged across three delays). Similarly, no significant relationship was found between averaged PM performance and averaged reminder usage.

4.3.4.5 Modification across Three Delays

First, we evaluated how metacognitive judgment modified reminder usage across different delays. A separate regression analysis was performed for each participant, looking at A) the relationship between the confidences (averaged across the two judgements) at the three delays, and B) the relationship between reminder usages at the three delays. For each group, a one-tailed Pearson correlation between the two resulting beta values showed a significantly negative relationship. In the hidden-clock group, participants whose confidence dropped more deeply at longer delays increased their reminder usage more sharply at longer delays ($r(46) = -.52, p < .001$). The association was just missed significance ($r(46) = -.21, p = .077$) in the persistent-on group.

Similar analyses were performed for each participant on A) the relationship between the reminder usage at the three delays, and B) the relationship between PAs at the three delays. For each group, a one-tailed Pearson correlation between the two resulting beta values was conducted, which showed no significantly negative relationship (persistent-on group: $p = .896$; hidden-clock group: $p = .063$). Therefore, participants who sharply increased reminder usage at longer delays did not show deeply falling performance at longer delays.

Then, the analyses were likewise performed for each participant on A) the

relationship between the confidences (averaged across the two judgements) at the three delays, and B) the relationship between PAWs at the three delays. For each group, a one-tailed Pearson correlation between the two resulting beta values was conducted, which showed no significantly negative relationship (persistent-on group: $p = .367$; hidden-clock group: $p = .722$). Therefore, participants whose performance dropped more deeply at longer delays did not have sharply decreased confidence at longer delays.

4.3.4.6 PM Performance Improvement

In the next analysis, we assessed the effect of offloading by comparing PM performance on both the unaided and reminder trials (See Figure 4.3 Middle Row). Specifically, repeated-measures $2 \times 3 \times 2$ ANOVAs on the measures of PM performance (i.e. PA, false hits, missing hits) were performed with within-subject factors Offload (unaided, reminder) and Delay (10s, 20s, and 30s) as well as a between-subject factor Clock (persistent-on, hidden-clock). The analysis revealed that the main effect of Offload and Delay were significant (Offload: $F(1, 94) = 99.76, p < .001, \eta^2_p = 0.52$; Delay: $F(1.7, 159.82) = 24.1, p < .001, \eta^2_p = 0.2$). However, the main effect of Clock was not significant. There was also a significant interaction between Offload and Delay ($F(2, 102) = 3.14, p < .001, \eta^2_p = .06$). As can be seen in Figure 4.3, the effect of offloading on the PM accuracy was more prominent at longer delay in both groups. The other interactions were not significant (Clock \times Offload: $F(1, 94) = 3.03, p = .08, \eta^2_p = .03$; Delay \times Clock: $F(1.7, 159.82) = .19, p = .79, \eta^2_p = .002$; Delay \times Clock \times Offload: $F(2, 188) = 1.06, p = .35, \eta^2_p = .01$). Further analyses showed the Clock \times Offload interaction was significant at 30s ($F(1, 94) = 4.22, p = .043$). Similar ANOVA analyses conducted on false hits and missing hits showed the main effect of Offload was significant (false hits: $F(1, 94) = 7.32, p = .008, \eta^2_p = .07$; missing hits: $F(1, 94) =$

31.6, $p < .001$, $\eta^2_p = .25$). Thus, the results overall demonstrated the effect of intention offloading, and PM performance was improved more profoundly at longer delays, which is congruent with the findings of Experiment 3a. Moreover, the hidden-clock group showed more prominent reminder effect on PM performance than the persistent-on group, but only at longer delays.

4.3.4.7 Ongoing Task Performance

Finally, we investigated whether the performance of the ongoing task differed among the three delay lengths, and was affected by reminder usage in the two groups (see Figure 4.3 The second row). A repeated-measures ANOVA analysing the d' measure from the 2-back task with within-subject factors, Offload (unaided, reminder) and Delay (no-target, 10s, 20s, 30s), as well as a between-subject factor, Clock (persistent-on, hidden-clock) revealed the main effects of Clock, Delay and Offload were all significant (Clock: $F(1, 94) = 15.29$, $p < .001$, $\eta^2_p = .14$; Delay: $F(2.53, 238.02) = 28.61$, $p < .001$, $\eta^2_p = .23$; Offload: $F(1, 94) = 58.66$, $p < .001$, $\eta^2_p = .38$). All the interactions were significant (Clock \times Delay: $F(2.53, 238.02) = 18.37$, $p < .001$, $\eta^2_p = .16$; Clock \times Offload: $F(1, 94) = 56.67$, $p < .001$, $\eta^2_p = .38$; Delay \times Offload: $F(2.09, 196.15) = 14.04$, $p < .001$, $\eta^2_p = .13$; Clock \times Offload \times Delay: $F(2.09, 196.15) = 4.64$, $p = .010$, $\eta^2_p = .047$). Post-hoc analyses showed the main effects of Delay and Offload were significant in the hidden-clock group (Delay: $F(2.37, 111) = 38.4$, $p < .001$, $\eta^2_p = .07$; Offload: $F(1, 47) = 68.4$, $p < .001$, $\eta^2_p = .2$), but not in the persistent-on group (Delay: $F(2.41, 113) = 2.27$, $p = .09$, $\eta^2_p = .003$; Offload: $F(1, 47) = .03$, $p = .87$, $\eta^2_p < .01$). The Delay \times Offload interaction was significant only in the hidden-clock group (hidden-clock: $F(1.98, 93) = 12.9$, $p < .001$, $\eta^2_p = .026$; persistent-on: $F(2.28, 107) = 1.96$, $p = .139$, $\eta^2_p < .01$). Thus, participants in the hidden-clock group had worse performance (i.e. d' -prime) than the persistent-on group. The d' of the

hidden-clock group was significantly improved when they could use reminders, and the improvement was greater at longer delays. However, in the persistent-on group, neither offloading nor delay length had a significant effect on the d-prime of the 2-back task.

We also analysed the RT measure with within-subject factors, Offload (unaided, reminder), Match (match, nonmatch), and Delay (no-target, 10s, 20s, 30s), as well as a between-subject factor, Clock (persistent-on, hidden-clock). There were significant main effects of Delay ($F(2.73, 256.53) = 13.04, p < .001, \eta^2_p = .12$) and Match ($F(1, 94) = 19.53, p < .001, \eta^2_p = .17$), but the main effects of Offload and Clock were not significant (Offload: $F(1, 94) = .25, p = .621, \eta^2_p < .01$; Clock: $F(1, 94) = .021, p = .886, \eta^2_p < .01$). The Delay \times Clock interaction was significant ($F(2.73, 256.53) = 3.09, p = .032, \eta^2_p = .03$). None of the other interactions were significant (Clock \times Offload: $F(1, 94) = 3.48, p = .065, \eta^2_p = .04$; Clock \times Match: $F(1, 94) = .02, p = .903, \eta^2_p < .01$; Delay \times Offload interaction ($F(2.56, 240.79) = 2.05, p = .117, \eta^2_p = .02$); Offload \times Match interaction ($F(1, 94) = .73, p = .974, \eta^2_p < .01$), Delay \times Match interaction ($F(2.64, 247.71) = 1.973, p = .127, \eta^2_p = .02$). Response times were significantly shorter on the match trials compared with nonmatch trials (hidden-clock: match: $M = 660\text{ms}, SD = 222$; nonmatch: $M = 702\text{ms}, SD = 197$; persistent-on: match: $M = 656\text{ms}, SD = 199$; nonmatch: $M = 696\text{ms}, SD = 161$). In the persistent-on group, participants having no PM target responded to the ongoing task more quickly (no-target: $M = 659, SD = 177$; 10s: $M = 671, SD = 186$; 20s: $M = 683, SD = 184$; 30s: $M = 690, SD = 181$) while no such difference was noted in the hidden-clock group. In conclusion, participants in both groups did not respond more rapidly to the 2-back task when they could set reminders in the PM task. The response time was shorter in the target trials compared with the non-target trials of 2-back task.

Table 4.2.

Average scores of the PM measures for each delay length and analyses of variance examining the effects of Clock, Delay and the interaction on the measure scores (F-values, significance levels and partial eta-square). Standard errors of mean are shown in parentheses.

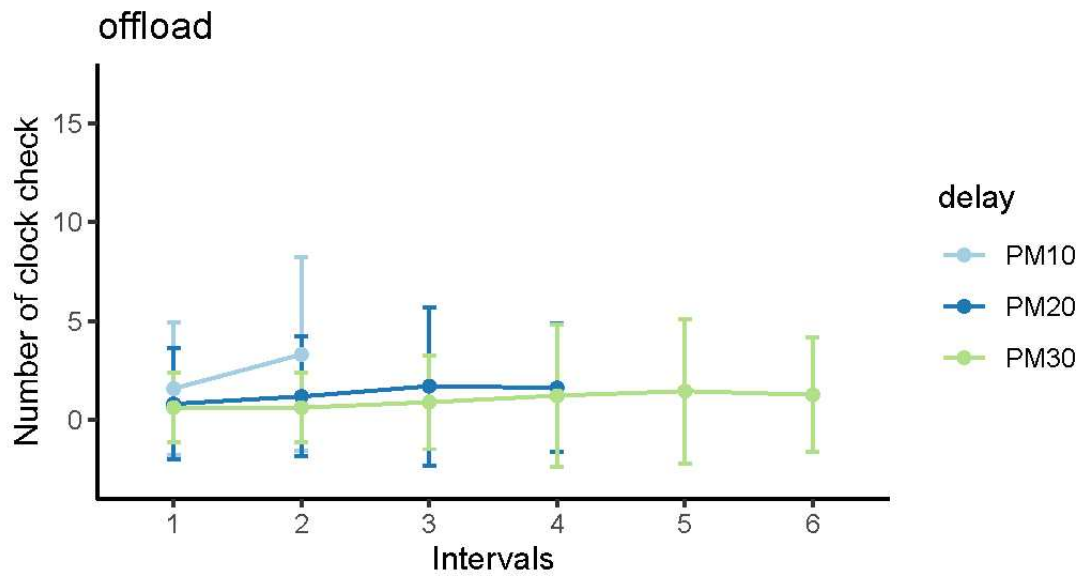
Measure	Hidden-clock group			Persistent-on group			Clock		Delay		Clock × Delay	
	10s	20s	30s	10s	20s	30s	F	η^2_p	F	η^2_p	F	η^2_p
PA (%)	85.6 (12.5)	76.0 (18.6)	70.0 (23.0)	84.5 (16.3)	75.5 (19.3)	73.3 (21.9)	.03	< .01	25.01 ***	.21	.74	< .01
False hits	1.69 (2.23)	3.75 (4.90)	4.96 (5.68)	3.52 (17.7)	2.54 (3.33)	10.2 (40.4)	.56	< .01	1.55	.02	.54	< .01
Missing hits (%)	.42 (.71)	.6 (.87)	.58 (.92)	.92 (1.11)	1.17 (1.53)	.96 (1.43)	7.66 **	.08	1.35	.01	.26	< .01
Confidence (%)	69.8 (18.7)	56.5 (21.8)	52.7 (24.7)	75.9 (16.9)	71 (15.5)	65.4 (20.6)	9.18 **	.09	42.73 ***	.31	4.25 **	.04
Metacognitive bias (%)	-15.8 (20.6)	-19.5 (22.2)	-17.3 (26.1)	-8.63 (21.1)	-4.48 (21.4)	-7.83 (25.6)	7.69 **	.08	.03	< .01	1.47	.02
Reminder usage (%)	65.6 (40.3)	80.9 (30.8)	87.7 (22.8)	38 (41.3)	58 (37.3)	62.5 (37.1)	14.15 **	.13	51.93 ***	.36	.48	< .01

Note: The calculation of the measures except reminder usage only includes the PM trials in the unaided conditio

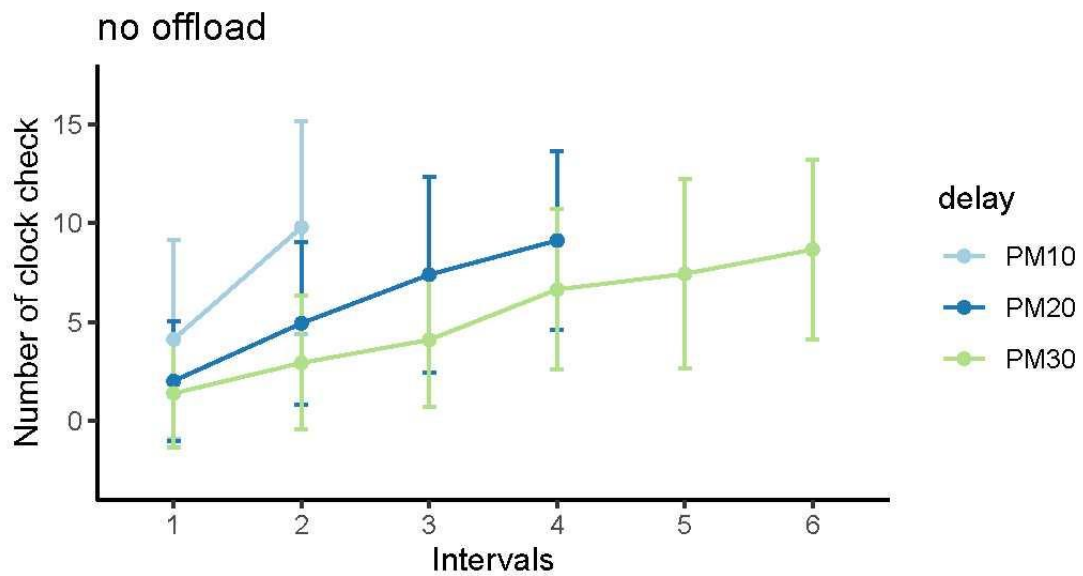
Figure 4.4

Mean number of clock checks prior to the designated time of the PM task in the hidden-clock group

A



B



Note.

PM10 (i.e. PM trials with a 10s delay): The number of clock checks averaged across subjects was counted in two interval, interval 1 (from 0s to 5s), and interval 2 (from 5s to 10s).

PM20 (i.e. PM trials with a 20s delay): The number of clock checks was counted in four intervals, interval 1 (from 0s to 5s), interval 2 (from 5s to 10s), interval 3 (from 10s to 15s), and interval 4 (from 15s to 20s).

PM30 (i.e. PM trials with a 30s delay): The number of clock checks was counted in six intervals, interval 1 (from 0s to 5s), interval 2 (from 5s to 10s), interval 3 (from 10s to 15s), interval 4 (from 15s to 20s), interval 5 (from 20s to 25s), and interval 6 (from 25s to 30s).

Error bars represent between-subject standard deviation.

A. offload: the PM trials where participants were allowed to set reminders

B. no offload: the PM trials where participants were forced to use unaided memory

4.3.4.8 Additional, Non-preregistered Analyses

This section reports some exploratory tests conducted in addition to the pre-registered ones described above. First, we conducted a one-tailed one-sample t-test on the metacognitive bias (averaged across the delay lengths). The results showed in both groups, participants were significantly underconfident of the unaided memory performance (persistent-on: $t(47) = -6.98, p < .001$; hidden-clock: $t(47) = -17.53, p < .001$).

Second, we conducted a one-tailed Pearson correlation showing confidence (averaged across delays) was correlated with unaided PM accuracy (averaged across delays) significantly in both persistent-on ($r(46) = .34, p = .019$) and hidden-clock groups ($r(46) = .45, p = .001$). Therefore, participants generally had reasonable, but not entirely precise, insight about how well they performed in the time-based PM task. This might raise a question about the contradicting results since the prior result showed participants were significantly underconfident of unaided memory performance. However, although people might generally underestimate their unaided memory performance, those who has worse unaided memory performance would exhibit lower confidence in their memory ability.

Third, we used lmer (Kuznetsova et al., 2017) to perform a linear mixed analysis evaluating the relationship between metacognitive judgement and reminder usage across different delays. As fixed effects, the model had variables: a) delay length (i.e. 10, 20, and 30 s) b) confidence (averaged across the two judgements) c) clock revealability (coded as persistent-on = 1, hidden-clock = -1) d) PM accuracy. As random effects, there were intercepts for participants. Likelihood ratio tests were conducted to compare the full model with the effect in question against the model without the effect in question (confidence: $\chi^2(1) = 12.64, p < .001$, PM accuracy: $\chi^2(1) = .16, p = .69$, clock revealability: $\chi^2(1) = 8.98, p = .003$). The results suggested confidence and clock revealability contributed to unique variance of reminder usage (confidence: $\beta = -.35, SE = .09$, clock revealability: $\beta = -10.65, SE = 3.50$).

Finally, as shown in Figure 4.4, the number of time-monitoring in the time interval closest to the designated time was greatest, which meant people checked the clock most frequently before the time. We conducted a Pearson correlation and found time monitoring frequency (averaged across delays) was correlated positively with unaided PM accuracy (averaged across delays) in the hidden-clock group ($r(46) = .43, p = .002$). These findings were in line with the prior research (Mioni 2014, Minoni 2019) indicating the relationship between the PM performance and the frequency of time monitoring behaviour. On the other hand, the pattern of the highest frequency for the latest interval was not observed when people were allowed to set reminders. It is likely that people were waiting for the clock to flash.

4.4 Discussion

The current experiments investigated a relatively unexplored issue, factors

influencing intention offloading in time-based PM tasks. Seeing as time acts as an important element of time-based PM, the effect of time-related factors could be distinct in a time-based PM task. First, participant's PM performance declined at longer delays, and the performance reduction was compensated by offloading intentions more often. Critically, we found that participants modified intention offloading on the basis of their metacognitive judgements about unaided memory ability. Second, participants increased intention offloading in the PM task where the clock revealability was manipulated to increase the cognitive load. Thus, participants tended to use more reminders in a time-based PM task if: a) they believed longer delay duration predicted worse performance, b) at the higher cognitive load condition where the clock was not persistently displayed.

Participants overall were underconfident about their performance in the time-based PM tasks, possibly as a result of the challenging nature of the current PM and ongoing tasks. However, people displayed fair metacognitive judgement about their performance. This is reflected in the significant correlations between the averaged confidence (averaged across delays) and averaged PM performance (averaged across delays). Some previous event-based PM studies (Gilbert 2015b; Gilbert 2020, Experiment 1, and Experiment 2) showed there was a relationship between reminder usage and objective performance. However, the current results along with the study of Boldt and Gilbert (2019) demonstrated the notion that people's metacognitive judgment predicts the extent to which people offload intentions while their objective unaided accuracy does not. First, the linear mixed-level analysis showed confidence predicted unique variance of reminder usage. PM performance was, however, not a significant predictor. Moreover, participants who sharply increased the reminder setting at longer delays had a greater drop of confidence at longer delays, but their

PM performance did not fall more deeply at longer delays. Thus, not only individuals who were less confident of unaided memory would set more reminders, but also, for each individual, metacognitive judgment guided the extent to which intention offloading was adjusted across different delays. The effect of metacognitive judgment was present irrespective of people's actual performance.

This suggests that metacognitive judgment may be a better predictor of reminder usage than memory performance. The study did not find a significant association between each individual's metacognitive judgements across three delays and memory performance across three delays. One possible explanation is that although people might have reasonable insight into their general memory performance in three delay conditions, they might have less aware of performance changes across different delay conditions. Further research could be conducted to evaluate this possibility.

4.4.1 Time Monitoring Behaviour

It has been suggested the management of clock revealability manipulates the cognitive load of the PM task. The hidden clock group considered as the group under higher cognitive load, however, did not have worse unaided PM performance. However the hidden clock group had worse performance (i.e. d-prime) on the ongoing task than the persistent-on group. Given previous research (Joly-Burra et al., 2018; R. E. Smith et al., 2012) has revealed the performance of the ongoing task was reduced when participants were given a PM instruction, and the performance reduction was considered as the costs of the PM task, it is possible that people exerted less cognitive efforts for the ongoing task to free cognitive resources for a more difficult PM task.

Regarding the main question whether clock revealability influences reminder usage, our results were in line with prior event-based PM studies (Gilbert, 2015a; Scarampi & Gilbert, 2021) which have suggested that the higher cognitive load could lead people to set more reminders. As expected, participants in the hidden-clock group used more reminders compared with the persistent-on group. The lower confidence of the hidden-clock group compared with the persistent-on group might contribute to the increase of reminder usage. On the other hand, seeing as the linear mixed analysis showing clock revealability contributed unique variance of reminder usage, it is plausible to assume that individuals in the hidden-clock group might set more reminders if the individuals had the same confidence as the individuals in the persistent-on group. One possible explanation is people could be biased to offloading in more demanding tasks. Setting reminders not only reduces cognitive load by offloading target-related contents onto the reminders, but also requires additional cognitive or physical effort (e.g. switch attention to the reminder button or press the button). Considering both the benefits and costs, it is more likely that people rely on internal memory for easier tasks, and use more reminders when the task requires more cognitive effort.

In terms of metacognitive bias, participants overall had negative bias and the hidden group had greater underconfidence than the persistent-on group. Similar findings were noted in Scarampi and Gilbert's event-based PM study (2021) where young and older participants performing an intention-offloading task were instructed to remember 1 target in half of the trials and 3 targets in the other half. The experiment included two phases. While in one phase, participants could choose to rely on reminders, they were forced to use the internal memory in the other phase. Young participants showed a greater negative metacognitive bias when there were 3

targets (compared with 1 target). It is not clear whether the increased cognitive load of PM tasks could increase people's metacognitive bias towards underconfidence. This would not reconcile with a well-known cognitive bias, the "hard-easy effect" (Lichtenstein & Fischhoff, 1977), which suggests people are overconfident at hard tasks and underconfident at easy ones. A possible explanation is that people might be biased to underconfident of the unaided memory ability when they are allowed to offload intentions rather than relying on the internal memory at a more difficult task. The argument provides a possible theoretical account that could be evaluated fully in the future.

4.4.2 Reminder Benefit

PM performance was improved significantly when participants were allowed to offload intentions. The benefit of reminders was greater at longer delays. This can be explained by the fact that participants used more reminders at longer delays. The difference in the benefit of reminders was, however, not found between the two clock groups though participants in the hidden-clock group used reminders more often. As well as the PM performance, another possible manifestation of reminder benefit is the performance of the ongoing task. No study as we know so far has examined the benefit of reminders for the ongoing task performance in a PM experimental paradigm. The current study showed the d-prime was improved by using reminders in the hidden-clock group, and the benefit was greater at longer delays. In contrast, the benefit for the d-prime was not significant in the persistent-on group. No significant benefit of reminders for the reaction times was noted across all delays and groups. Thus, the intricate pattern of the reminder benefit could offer a future research avenue that the diverse measures for assessing the benefit of reminders should be developed in either laboratory or naturalistic experiments.

4.4.3 Limitation and Future Research

The main methodological limitation of the present study was the brief delay duration used in the paradigm. This study observed a ceiling effect in the time-based PM task, which may be attributed to this unique feature that deviated from traditional TBPM paradigm which usually employed a longer delay. Therefore, it may be challenging to compare these results to previous time-based literature. However, the use of short delays allowed for a larger number of trials, which facilitated sufficient data collection. The current study along with the findings of Einstein et al. (2003) demonstrated maintaining an intention over a short time such as less than 1 minute still required moderate cognitive resource. It was reflected in the PM performance decline which was affected by experimental manipulations such as delay length or divided attention. However, people usually need to maintain intentions for a longer time in everyday time-based PM tasks. It remains uncertain people would offload intentions likewise in everyday time-based PM tasks. A promising approach in the future is to evaluate whether people in naturalistic settings would adjust intention offloading in a time-based task on the basis of metacognitive judgment.

The current research on delayed intention and time-based PM could broaden the knowledge about the cognitive process modifying the decision about offloading intentions. Given the effect of metacognitive judgment on the reminder usage has been consistently found in the previous event-based PM literature as well as the current study, it might be plausible to indicate that training the metacognitive judgment improves the application of offloading strategies for PM tasks independently of the task type (i.e. event-based and time-based). In light of the present findings, there are two possible future avenues for the metacognition-

related intervention. First, the metacognitive training is possibly more helpful if the intervention could include not only the common conditions, but also the knowledge regarding the extent to which the optimal usage of reminders could be influenced by some factors such as delay duration, cognitive load. People might have decent metacognitive judgment about general unaided ability in a PM task, but relatively incompetent metacognitive judgment about the differences in the performance when the PM task are altered partially. Second, solely giving the metacognitive advice about PM performance might fail to manifest how individual manage the whole cognitive resource on both PM and ongoing tasks, and mislead people to set reminders suboptimally. In the present study, participants had similar PM performance irrespective of the clock revealability, which means the metacognitive advice about PM performance would be alike. However, setting more reminders in the PM task allowed participants in the hidden-clock group to better perform the ongoing task.

Another issue raised by our results concerns the modification of intention offloading is not only triggered by metacognitive judgment, but also cognitive load. People might be biased towards offloading in a more demanding task regardless of how confident they were. In the event-based PM literature, there are some evidence showing people were biased towards offloading to avoid cognitive effort (Gilbert et al., 2020; Sachdeva & Gilbert, 2020). A further approach could evaluate whether the extent to which people are bias towards offloading is impacted by cognitive demand of the task when the metacognitive judgment is controlled.

4.4.4 Conclusion

In summary, along with the prior two experiments, the current study showed that people adjust intention offloading based on metacognitive judgment regardless of event-based or time-based. Two crucial time-associated factors, delay length and time-monitoring, in a time-based PM task impact the extent to which people chose to offload. The impact could not be fully explained by metacognitive judgment so other factors such as cognitive demand should be considered.

Experiment 4: The Neural Correlates of Cognitive offloading: Compared with Intentional Forgetting and Remembering

5.1 Introduction

In Experiment 1, 2, 3, we investigate various factors influencing cognitive offloading. Following this, Experiment 4 evaluated the brain correlates following the occurrence of cognitive offloading. As discussed in Chapter 1, Storm and Stone(2015) argue that list-method directed forgetting (LMDF), and saving information could lead to pre-cue memory reduction and post-cue memory enhancement, which implies a correspondence between cognitive processes involved in the two phenomena, directed forgetting and information offloading.

Extensive literature has suggested that directed forgetting involves an active process triggering the forgetting by given instructions. The active process could interrupt mnemonic function during the encoding and retrieval phases (Johnson, 1994). During the encoding phase of the second list, prior LMDF studies found that increased activation in the prefrontal area and reduced neural synchrony was correlated with directed forgetting (Bäuml et al., 2008; Hanslmayr et al., 2012). Hanslmayer et al. (2012, Experiment 1), using list-method paradigm and fMRI scans, found that people in the forget trials who were cued to forget List 1 had increased brain activity in the left dorsolateral prefrontal cortex (DLPFC) and decreased DLPFC-hippocampal synchrony during List 2 encoding. Furthermore, the study (Experiment 2) revealed through using transcranial magnetic stimulation to stimulate the left

DLPFC during List 2 encoding that directed forgetting was enhanced and DLPFC-hippocampal synchrony was reduced. Manning et al. (2016) found by using pattern classifiers that forget cues affect the activation of the contextual representations of the first list to a lower extent than the remember cues, and the reduced recall proportion of the first list in forget trials was associated with this lower extent of activation. Anderson and Hulbert (2021) suggested that the reduced DLPFC-hippocampal synchrony during the forget condition supported the context change account of directed forgetting: the prefrontal control mechanism alters the representations of mental context and remove cues required for memory access.

On the other hand, retrieval inhibition has been proposed as a possible mechanism of directed forgetting. To investigate the inhibition process that occurs during the retrieval phase, Anderson and Green (2001) developed the think/no-think (TNT) paradigm in order to investigate how suppressing unwanted memories have been facilitated by retrieval inhibition. In the TNT paradigm, participants were first required to study cue-target pairs such as “work cat” or two pictures. Then, they were instructed to recall the target when a cue was presented in a specific colour (i.e., think trials), but suppress the retrieval of the target when a cue was in another colour (i.e., no-think trials). In the final test phase, participants recalled the target when cues were given. Through comparing the memory performance between think, and no-think trials, the TNT paradigm shows whether people engage in retrieval inhibition. Compared with the think condition, the no-think condition yielded signal change in two sets of brain regions. One set of regions, that is, right-lateralised regions, including the right prefrontal cortices, posterior middle frontal gyrus and bilateral insula (Anderson et al., 2004, 2016; Depue et al., 2007; Schmitz et al., 2017), increased their activities when participants stopped retrieving the memory. A second

set of regions decreased their activities during retrieval inhibition, including the hippocampus (Depue et al., 2007; Levy & Anderson, 2012) along with relevant cortical and subcortical regions engaging in the processes of encoding information (Gagnepain et al., 2014; Mary et al., 2020). For example, hippocampus triggers reinstated activities in the fusiform gyrus during the retrieval of visual objects, and the propagating activity in both the hippocampus and fusiform gyrus was suppressed when people inhibited retrieving the memory (Gagnepain et al., 2014; Mary et al., 2020). Between these two sets of regions, previous connectivity analysis research has shown that there is a causal coupling, a frontal-hippocampal inhibitory control, occurring during retrieval inhibition (Benoit et al., 2015; Cook et al., 2007; Gagnepain et al., 2014, 2017).

In spite of the increasing body of research that has focused on better understanding neural correlates supporting directed forgetting during the encoding and retrieval phase, few studies have evaluated the brain activity occurring at the moment of Forget cue displaying. An item-method directed forgetting (IMDF) study (Oehrns et al., 2018) using intracranial EEG recording found increased theta oscillations at dorsolateral prefrontal cortices and oscillations in the alpha/beta frequency band at the hippocampus at 568-1,058 ms after the onset of Forget cues. This indicates that the prefrontal inhibition of hippocampal processing might support directed forgetting. On the other hand, a recent IMDF study (Wang et al., 2019) proposing an alternative mechanism of directed forgetting found that the moderate activation of ventral temporal cortex in the instruction phase was associated with directed forgetting. The results are consistent with nonmonotonic plasticity hypothesis, suggesting that while higher level of memory activation enhances memory, activating memory at a moderate level renders the ability to recall to a

reduced state (Detre et al., 2013; Lewis-Peacock & Norman, 2014; Newman & Norman, 2010; Norman et al., 2007). Although the IMDF studies have revealed possible neural correlates of directed forgetting in the instruction phase, it remains a question whether similar brain activation pattern would emerge in a LMDF paradigm.

The background presented so far lead up to the question: would offloading information have similar mechanisms as directed forgetting? Runge et al. (2021) used Storm and Stone (2015) list-method paradigm to evaluate the EEG oscillations during the second list encoding and compared the neural correlates of saving effects to the findings of a previous LMDF study (Hanslmayer et al., 2012), which showed that a reduction of alpha power and alpha phase synchrony during List 2 encoding were related to the memory enhancement effect of directed forgetting. However, the results neither replicated the saving-enhanced memory effect (Storm and Stone, 2015, see Chapter 1), nor the reduction of alpha power and alpha phase synchrony. To explain the results, the study proposed that saving-enhanced memory effect might be smaller than presumed. Further studies are therefore needed to clarify the answer to the question whether saving is different from directed forgetting and what neural correlates could characterise the difference.

In the current study, we used fMRI to investigate the following topics:

- 1) neural correlate of a remember, forget, or offload cue
- 2) behavioural effects of List 1 cue on List 2 memory

The overall paradigm design, similar to the LMDF approaches and the procedure used by Storm and Stone (2015, see Chapter 1) in Experiment 1, was adapted for an fMRI environment with several changes. First, participants were presented with images of faces and scenes. Each item in a list was presented one by one in order

rather than the whole list at once. After the first list, participants were instructed to forget, remember, or offload by a cue. The forget cue instructed participants that they would not be tested on the items in the first list. The remember cue instructed participants that they might be tested on the items in the first list. The offload cue instructed participants that they might be tested on the items in the first list, but they would receive a reminder if so. Participants were not actively setting reminders but were passively receiving the answers. This approach was because tasks such as actively setting reminders might interfere with brain activities, and complicate the interpretation of fMRI findings. At the end of each trial, participants were shown one of the items and required to indicate its position (first, second, third etc.) within its list. Memory was tested on one of the two lists, not both lists. Having finished the MRI scan, participants received a surprise test, a recognition test, outside the scanner of all the items displayed in the experiment, including the items which were instructed to forget. This examination allows us to investigate whether the effects of directed forgetting or offloading would be temporary or sustain over time.

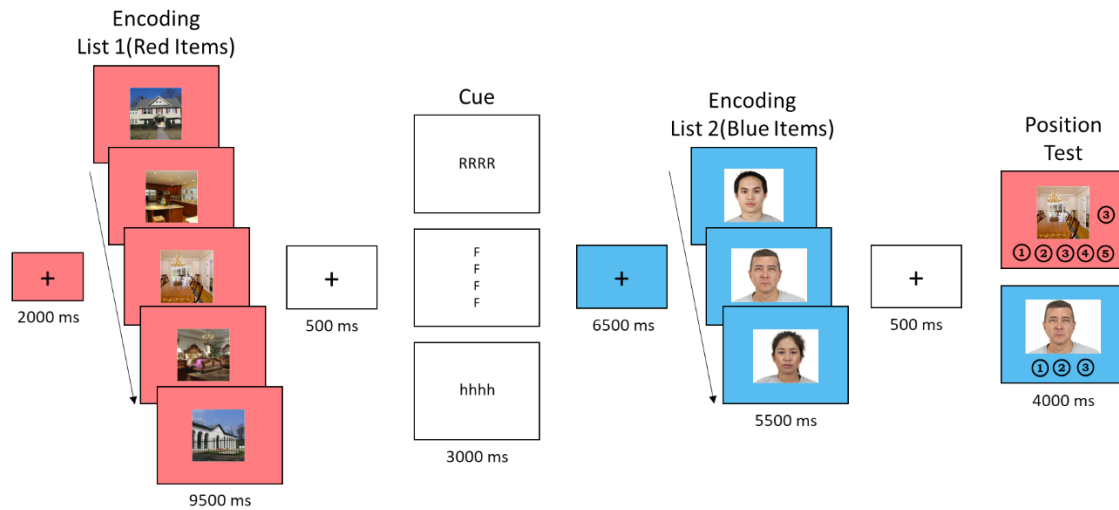
5.2 Method and Procedure

5.2.1 Stimuli

Experimental stimuli include colour images of scenes and faces. A large collection of face stimuli from the Chicago Face Database (CFD), CFD-MR extension and CFD-INDIA extension was used in the present experiment (Lakshmi et al., 2021; Ma et al., 2015, 2020). Faces had neutral expressions, were cropped from the neck down, and shown over a white background. A subset of scenes from the Fine-Grained Image Memorability Dataset was also used in the present experiment (Bylinskii et al., 2015). All items were sized to 300 × 300 pixels and presented using

Figure 5.1

A Schematic Illustration of the List-method Directed Forgetting and Saving Task



Note. Participants performed list-method directed forgetting and saving task on faces and scenes in the scanner. During each trial, participants first viewed five List 1 images sequentially (i.e. Red items). Then they received a memory cue instructing them to remember, forget or offload List 1 images. After the cue, they viewed three List 2 images (i.e. Blue items) sequentially. At the end of each trial, they were given a position test for one of the two lists. Participants were shown one of the items and required to indicate its position (first, second, third etc.) within its list. In the offload condition (i.e. help condition), participants might be tested on the items in List 1 (Red items), but they would receive a reminder.

5.2.2 Method

The list-method directed forgetting and saving (LMDFS) task comprised 7 blocks, each including 12 trials. The schematic illustration of the procedure on a given trial is shown in Figure 5.1. Subjects were allowed to have a break between blocks. Each trial involved the encoding and testing of the contents of two lists, where each list either consisted of a sequence of face images or a sequences of scene images.

Subjects were first shown the first list, either five faces or scenes in sequence, each for 1.5 s, which were separated by a 0.5 s blank screen. The stimuli of the first list were called red items because the background colour was red. Following the presentation of the red items, an instruction cue was given for 3 s. The cues for forget, remember, and help trials were represented by the letters f, r, and h respectively. Four letters, either uppercase or lowercase, were aligned horizontally or vertically in the center of the screen. There were two possible cues for each instruction: half of the participants had half uppercase horizontal cues and half lowercase vertical cues, while the other half of the participants had half uppercase vertical cues and half lowercase horizontal cues. Participants were instructed to apply the instruction represented by the cue to the first list. Participants were not suggested to use any particular strategy. On the contrary, they were instructed to simply forget or remember the previously presented red items in the forget or remember trials. In the help trials, they were instructed that the answer would be displayed if the red items were tested at the end. After a post-cue screen with fixation cross for 6.5s, subjects were shown the second list, either three faces or scenes in sequence, each for 1.5 s, which were separated by a 0.5 s blank screen. The images of the second list were called blue items because the background color was blue. The length of two lists was based on our pilot studies, which showed that saving effect was more prominent in this experimental paradigm. At the end of each trial, there was a position test where an image, either one of the red items or the blue items, was displayed and participants would provide as an answer the order of the image in the list (e.g., the “second” image of the first list) by pressing one of the buttons representing “First,” “Second,” “Third,” “Fourth,” or “Fifth.” For the forget trials, there was no position test in half of the trials, but in the other half of the trials, the images in the second list was tested. This meant that the images from the second

list were equally likely to be tested, regardless of the earlier cue. Half of the remember trials were tested on the first list while the other half were tested on the second list. Half of the help trials were tested on the first list, and the image was displayed with an answer (e.g. "2") so that the subject did not need to recall the order while the other half of the help trials were tested on the second list. Which image in a list was tested in the position test was chosen randomly. One-third of the trials were forget trials, one-third of the trials were remember trials, and the remaining trials were help trials. These were presented in a randomised order. The categories, and categorical compatibility for List 1 and List 2 (i.e. face or scene) were counterbalanced over the 12 trials (i.e. each block): 1) for 3 trials, List 1 items were faces and List 2 items were scenes; 2) for 3 trials, List 1 items were faces and List 2 items were faces; 3) for 3 trials, List 1 items were scenes and List 2 items were scenes; 4) for 3 trials, List 1 items were scenes and List 2 items were faces.

5.2.3 Participants

Thirty-six healthy subjects between the ages of 18 and 34 were recruited from the University College London as well as from the surrounding community. Participants first completed the on-line phase. Then, we excluded participants who had lower accuracy in the reminder trials compared to the remember trials, as well as those whose accuracy rate was below 70% in the remember trials. The participants whose results fit the inclusion criteria (see the criteria below) were invited to join the second phase including a MRI scan and a post-MRI memory test. Participants were compensated at a rate of £7.5 for the one-hour on-line phase and £20 for the MRI phase. Informed consent was obtained from all subjects. All subjects were right-handed and had normal or corrected-to-normal vision. The study was approved by the UCL Research Ethics Committee (1584/002), and informed consent

was obtained from all participants.

A total of 36 subjects (31 female; mean, 23.3 years old) are included in the reported fMRI analyses. For three of these subjects, a malfunction of online data storage affected the behavioural data of their recognition memory test after the fMRI scans, but these three subjects completed the LMDFS task in the scanner, contributed the position test data, and were included in the position test and fMRI data analyses.

Inclusion Criteria

1. Accuracy of help trials above 70%
2. Higher accuracy of help trials than remember trials

5.2.4 Procedures

The current experiment included two phases. Participants first completed the LMDFS task online, and then the participants who were not excluded joined the second phase comprising of a MRI scan and a post-MRI memory test.

5.2.4.1 On-line phase

The task was programmed using Gorilla (<https://gorilla.sc/>). After providing informed consent, participants engaged in practice trials, explaining the different features of the tasks one at the time. Participants were then introduced to the LMDFS task consists of 7 blocks, each including 12 trials. At the end of the experiment, participants were thanked for their time, paid, and debriefed.

5.2.4.2 MRI Scan

The items shown in the task of the on-line phase were not used for the second

phase. In this phase, each subject completed two tasks: the LMDFS task and the post-MRI memory test. Before the LMDFS task, subjects first did a few practice trials to refresh their memory for the task. The LMDFS task was administered in the MRI scanner, while the memory test was administered outside the scanner, immediately after the LMDFS task was completed.

5.2.4.3 Post-MRI Memory Test

The post-MRI memory test was a self-paced recognition test conducted outside the scanner. Participants performed a recognition memory test on a large set of 336 items, which included 168 items from the LMDFS task (half faces, half scenes), which were not tested in the position tests and 168 new items. They were tested on List 2 items first followed by List 1 items, including the items instructed to forget. Participants would answer their confidence judgments (i.e. “Surely Old,” “Likely Old,” “Likely New,” or “Surely new”) when each image was displayed. Confidence responses were assigned points because participants were expected to respond based on the actual memory of items, rather than the instruction cue of items, such as an item followed by a forget cue. Although it was possible that participants gave incorrect responses in the forget condition because they knew they were not supposed to remember the associated image, it was challenging to remember which image corresponded to which cue during the post-MRI memory test. Subjects were informed of the point system and told that they should maximise their points. The point system was as follows: for each old item, a “Likely New” response was penalized with 0.5 point, and a “Surely New” response was penalized with 1 point while a “Likely old” was awarded 0.5 point, and a “Surely Old” response was awarded 1 point. For each new item, a “Likely old” was penalized 0.5 point and “Surely old” was penalized 1 point while a “Likely new” response was awarded 0.5 point, and a

“Surely new” response was awarded 1 point. The total sum was reported to the subject at the end of the memory test.

5.2.5 fMRI Recording and Data Analysis

Scanning was performed using a Siemens Prisma 3 T MR head scanner. (TR = 1.14s, TE = 35ms, multi-band acceleration factor = 4). Each volumes was comprised of 56 axial slices (2mm thick, comprising 106 x 106 2mm x 2mm square pixels), oriented approximately to the AC-PC line. Following the 7 functional runs, a 6-minute MPAGE T1-weighted structural scan was performed.

We used SPM12 (Wellcome Department of Cognitive Neurology, London, UK; www.fil.ion.ucl.ac.uk/spm) implemented in MATLAB R2016b (MathWorks, Sherbom, MA) to perform image pre-processing and statistical analysis. Pre-processing consisted of realignment, normalisation into 3mm cubic voxels with 4th-degree b-spline interpolation (performed via the co-registered structural scan), and smoothing with a 8mm full-width half maximum Gaussian filter. fMRI data were analysed in an event-related manner. Variance in each time series was decomposed in a voxelwise general linear model with the following regressors: 1) presentation of the list 1 items in the face condition (10-second boxcar regressor), 2) presentation of the list 1 items in the scene condition (10-second boxcar regressor), 3) forget cue (delta function), 4) remember cue (delta function), 5) remind cue (delta function), 6) list 2 items in the face condition (6-second boxcar regressor), 7) list 2 items in the scene condition (6-second boxcar regressor). These regressors, plus an additional regressor representing the mean over scans and six additional regressors representing the realignment parameters calculated by SPM12 comprised the full model for each session. The data were high-pass filtered at 256 s. On a single-subject level, we first contrasted the

three conditions of interest (Forget, Remember, and Help) with pairwise comparisons. Contrast maps of beta values derived from these comparisons (Forget vs. Remember, Remember vs. Help, and Forget vs. Help) were then entered into second-level one sample t tests. Planned comparisons were conducted using one-tailed t tests (height threshold $p < .001$) along with an extent threshold calculated by SPM to achieve a whole-brain familywise error corrected threshold of $p < .05$.

5.3 Results

5.3.1 Behavioural Results

5.3.1.1 Results for Position Test

We evaluated whether participants could successfully remember the position of an image in the sequences of two lists. The accuracy rate of List 1 items was significantly higher in the help ($M = 95.4$, $SD = 10.5$) than in the remember ($M = 56.0$, $SD = 21.2$) trials ($t(35) = 12.05$, $p < .001$, $d = 2.36$). The reaction time was shorter in the help ($M = 1.6$, $SD = .4$) than in the remember ($M = 2.3$, $SD = .4$) trials ($t(35) = 10.18$, $p < .001$, $d = 1.70$). This demonstrated that reminders improved memory performance on List 1.

In terms of List 2, differences ($F(2,70) = .09$, $p = .914$, $\eta_p^2 < .01$) were not found in accuracy, depending on the instruction cue for List 1 (forget: $M = 79.4$, $SD = 18.1$; help: $M = 80.2$, $SD = 18.4$; remember: $M = 80.4$, $SD = 19.9$); differences ($F(2, 70) = .54$, $p = .58$, $\eta_p^2 = .02$) in the reaction time across cue types (forget: $M = 1641$ ms, $SD = 409$; help: $M = 1675$ ms, $SD = 432$; remember: $M = 1681$ ms, $SD = 348$) were also not found.

We then compared the accuracy of List 2 between the compatible (i.e. List 1 and List 2 from the same category) and incompatible trials (i.e. List 1 and List 2 from the different category) in a 2x3 ANOVA with factors Cue and Category (compatible: forget: $M = 79.1$, $SD = 20.0$; help: $M = 82.6$, $SD = 23.2$; remember: $M = 75.3$, $SD = 21.6$; incompatible: forget: $M = 78.7$, $SD = 22.4$; help: $M = 77.5$, $SD = 21.6$; remember: $M = 84.0$, $SD = 21.4$). Results are shown in Figure 5.2. There were no significant main effects of Cue ($F(2,70) = .09$, $p = .891$, $\eta_p^2 < .01$) and Category ($F(1,35) = .09$, $p = .578$, $\eta_p^2 < .01$), but a significant interaction ($F(1,67) = .09$, $p = .023$, $\eta_p^2 = .11$). The interaction shows the effect of Cue was affected by the Category factor. Then we used paired t-tests to reveal the effects of categorical compatibility. Results showed that when participants were required to remember List 1, the memory performance of List 2 was worse in the compatible trials than the incompatible trials ($t(35) = 2.53$, $p = .016$, $d = .42$). No significant effect of categorical compatibility was noted in help and forget trials (help: $t(35) = 1.66$, $p = .107$, $d = .28$ forget: $t(35) = .10$, $p = .922$, $d = .02$). Also, we used paired t-tests to reveal the effects of cue on performance for compatible and incompatible trials separately. In compatible trials, List 2 accuracy was higher following a H-cued List 1 than following a R-cued List 1 ($t(35) = 2.24$, $p = .094$, $d = .37$) whereas the accuracy of List 2 following H-cued List 1 was lower than List 2 following R-cued List 1 ($t(35) = 2.10$, $p = .013$, $d = .35$) in incompatible trials. Thus the saving-enhanced memory effect was found in compatible trials, but not in incompatible trials. No significant difference was found between forget and remember trials or between forget and help trials (compatible: forget vs. remember: $t(35) = .91$, $p = .369$, $d = .15$, forget vs. help: $t(35) = .88$, $p = .383$, $d = .15$; incompatible: forget vs. remember: $t(35) = 1.39$, $p = .174$, $d = .23$, forget vs. help: $t(35) = .49$, $p = .630$, $d = .08$).

We next analysed the reaction times of List 2 between the compatible and incompatible trials in a 2×3 ANOVA with factors Cue and Category (compatible: forget: $M = 1644\text{ms}$, $SD = 447$; help: $M = 1648\text{ms}$, $SD = 439$; remember: $M = 1753\text{ms}$, $SD = 425$; incompatible: forget: $M = 1651\text{ms}$, $SD = 430$; help: $M = 1707\text{ms}$, $SD = 467$; remember: $M = 1626\text{ms}$, $SD = 348$). Results are shown in Figure 5.2. Consistent with the findings of accuracy, the results of reaction times showed no significant main effects in Cue ($F(2,70) = .55$, $p = .582$, $\eta_p^2 < .02$) and Category ($F(1,35) = .33$, $p = .568$, $\eta_p^2 < .01$), but a significant interaction ($F(1,67) = 5.23$, $p = .013$, $\eta_p^2 = .13$). Then we used paired t-tests to reveal the effects of categorical compatibility. Results showed that the reaction times of List 2 following a remember cue was longer in the compatible trials than the incompatible trials ($t(35) = 2.30$, $p = .081$, $d = .38$). No significant effect of categorical compatibility was noted in help and forget trials (help: $t(35) = 1.45$, $p = .155$, $d = .24$; forget: $t(35) = .15$, $p = .881$, $d = .03$). Also, we used paired t-tests to reveal the effects of cue on reaction times for compatible and incompatible trials separately. No significant difference was found between the three conditions (compatible: forget vs. remember: $t(35) = .46$, $p = .646$, $d = .08$; forget vs. help: $t(35) = .09$, $p = .932$, $d = .01$; help vs. remember: $t(35) = 1.78$, $p = .083$, $d = .30$; incompatible: forget vs. remember: $t(35) = 1.39$, $p = .174$, $d = .23$; forget vs. help: $t(35) = 1.73$, $p = .093$, $d = .29$; help vs. remember: $t(35) = 1.62$, $p = .113$, $d = .27$).

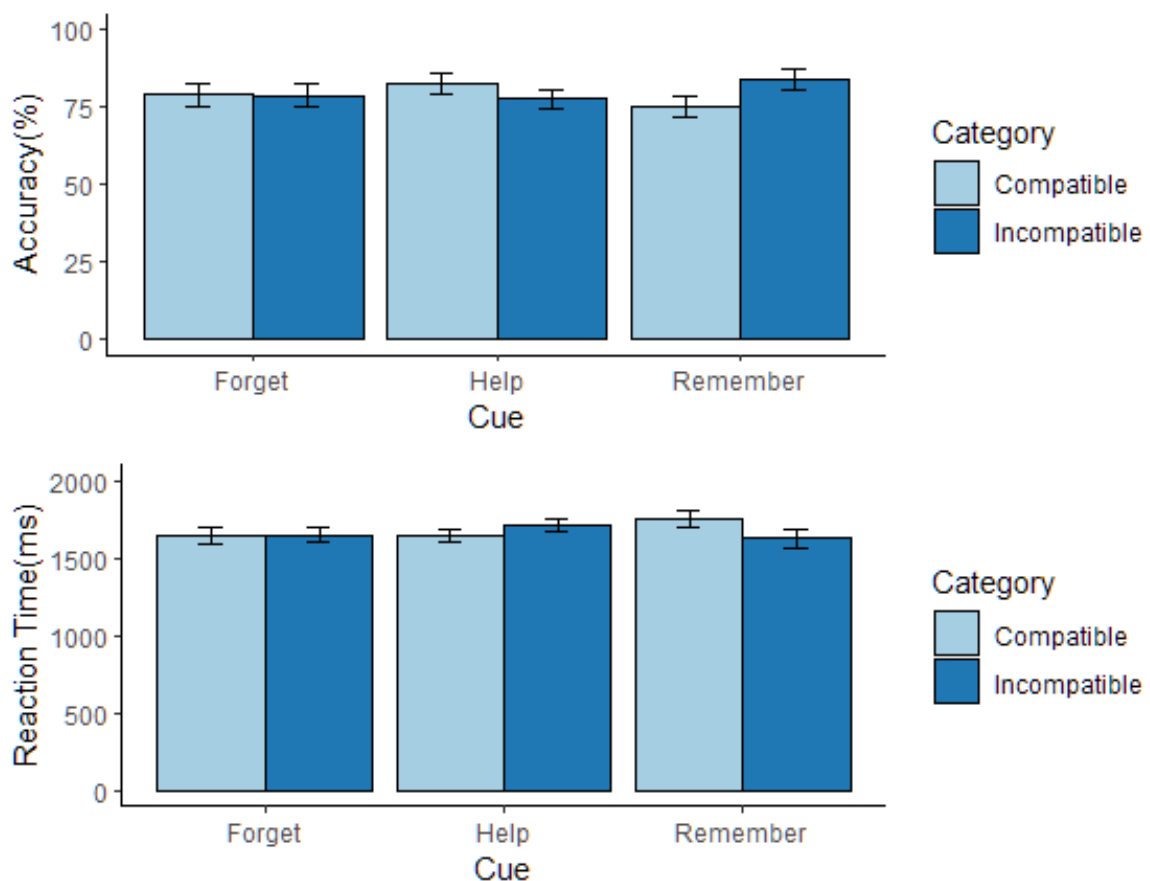
5.3.1.2 Results for Recognition Test

We examined the differences in the recognition rate on the images across three cue types after finishing the MRI scan. For both List 1 and List 2 items, no significant difference (List A: $F(2,64) = .68$, $p = .511$, $\eta_p^2 < .02$; List B: $F(2,64) = .03$, $p = .974$, $\eta_p^2 < .01$) in the recognition rate was found across three types of cues (List A: forget: $M = 79.4$, $SD = 18.1$; remember: $M = 80.4$, $SD = 19.9$; help: $M = 80.2$, $SD = 18.4$; List B:

forget: $M = 79.4$, $SD = 18.1$; remember: $M = 80.4$, $SD = 19.9$; help: $M = 80.2$, $SD = 18.4$); We used a 2×3 ANOVA with factors Cue and Category to compare the accuracy of List 2 between the same category and the different category trials in fMRI, and found no significant main effects or interaction (Cue: $F(2, 64) = .01$, $p = .989$, $\eta_p^2 < .01$; Category: $F(1,32) = .04$, $p = .842$, $\eta_p^2 < .01$; interaction: $F(2,64) = .47$, $p = .625$, $\eta_p^2 = .02$).

Figure 5.2

Behavioural Results from the Position Test: The Effects of Cues and Categorical Compatibility on List 2 Accuracy and Reaction Time of the Position Test



Note.

Upper row: List 2 accuracy of the position test as a function of Cue (forget, help and remember) and Category (compatible and incompatible). Error bars represent within-subject confidence intervals (comparing the compatible and incompatible conditions). The error bars do not overlap with each other, which means $p < .05$.

Lower row: List 2 reaction times of the position test as a function of Cue (forget, help and remember) and Category (compatible and incompatible). Error bars represent within-subject confidence intervals (comparing the compatible and incompatible conditions). The error bars do not overlap with each other, which means $p < .05$.

5.3.2 fMRI Results

Results are shown in Table 5.1. and Figure 5.3. We investigated the extent to which offload cues are functionally equivalent to forget cues, by comparing both conditions against the remember condition along with comparing them with each other.

5.3.2.1 Forget vs. Remember Trials

We examine brain activity associated with the presentation of the cue (Forget, Remember, or Help). When participants were presented F-cue vs. R-cue (F>R), widespread bilateral regions were activated with two peaks, one at right supramarginal gyrus ($k=12821$, MNI coordinates of peak voxel: 60, -37, 29), the other one at right precentral gyrus ($k=115$; 36, -16, 41), extending laterally to angular gyrus, inferior parietal gyrus, superior temporal gyrus, middle temporal gyrus, insula, medially to cingulate gyrus, parahippocampus gyrus, hippocampus gyrus, fusiform gyrus, lingual gyrus, thalamus, anteriorly to superior frontal gyrus and inferior frontal gyrus, posteriorly to partial occipital lobe. Overall, the right hemisphere activations appeared more extensive than the left hemisphere ones. Conversely, we observed decreased activation (R>F) in a few regions. One peak was located at the left inferior frontal gyrus ($k=369$; -51, 17, 32), extending to precentral gyrus and middle frontal gyrus. The other two peaks were located at bilateral inferior parietal gyrus (left: $k=314$; -33, -43, 41; right: $k=174$; 33, -52, 44), extending to the deep regions of angular gyrus, supramarginal gyrus and partial superior parietal gyrus.

5.3.2.2 Help vs. Remember trials

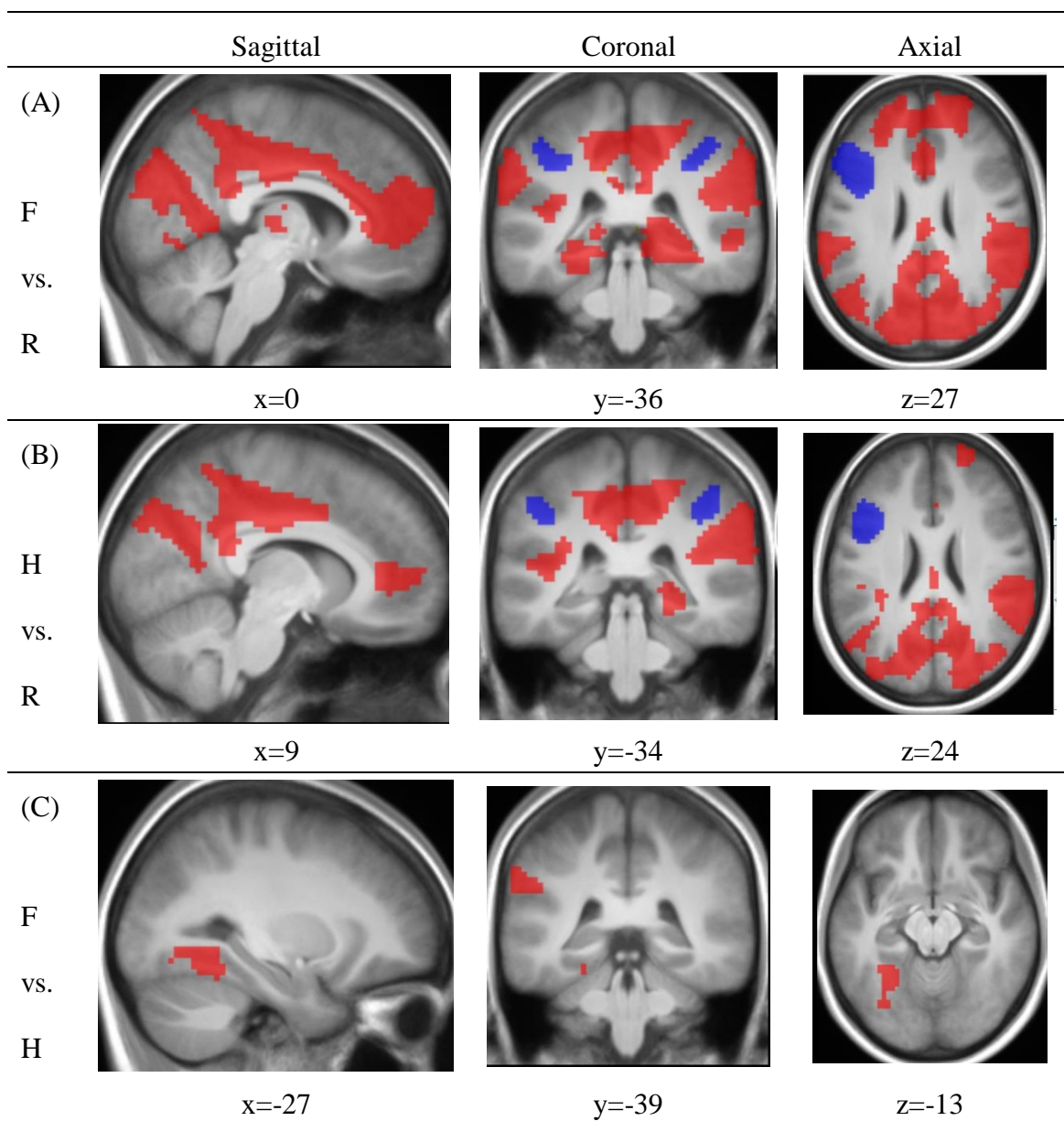
When participants were instructed by an H-cue vs. an R-cue (i.e. $H > R$), widespread activated regions were found. Three right peaks of the activation were located at paracentral gyrus ($k=6094$; 15, -31, 50), middle frontal gyrus ($k=519$; 30, 50, 35), and inferior frontal gyrus ($k=157$; 48, 29, 8), which extended laterally to angular gyrus, supramarginal gyrus, superior temporal gyrus, middle temporal gyrus, insula, medially to cingulate gyrus, parahippocampus gyrus, hippocampus gyrus, fusiform gyrus, posteriorly to partial occipital lobe. There was one peak in the left hemisphere which was located at planum polare ($k=504$; 42, -22, -1), spreading to angular gyrus, cingulate gyrus, fusiform gyrus and partial occipital lobe. On the other hand, we observed decreased activation ($R > H$) in a number of regions: a) deep regions of left inferior parietal gyrus, and supramarginal gyrus ($k=241$; -42, -43, 44) b) left precentral gyrus, and middle frontal gyrus ($k=330$; -45, 5, 35) c) caudate nucleus extending to partial insula ($k=103$; -15, 14, 8) d) deep regions of right supramarginal gyrus ($k=151$; 45, -31, 44).

5.3.2.3 Forget vs. Help trials

Comparing the brain activities elicited by a cue between the forget and help trials ($F > H$), we found higher activation at left supramarginal gyrus ($k=72$; -54, -37, 29) and left fusiform and lingual gyrus ($k=71$; -30, -58, -7). No significant decreased activation ($H > F$) was observed in this scenario.

Figure 5.3

Neural Correlates of directed forgetting, offloading and remembering



Note. Panels A-C display fMRI activations on sagittal, coronal, and axial slices respectively seen in the mean normalised structural scan ($p < .05$, FEW corrected for multiple comparisons at the cluster level with an underlying uncorrected threshold of $p < .001$).

- A. Forget cues contrasted with remember cues: red regions: F>R; blue regions: R>F
- B. Help cues contrasted with remember cues: red regions: H>R; blue regions: H>F
- C. Forget cues contrasted with help cues: red regions: F>H; no significant differences were observed in the H>F contrast.

Table 5.1*Results of the whole-brain analysis*

Anatomical label	MNI coordinates					
	HS	Size	<i>x</i>	<i>y</i>	<i>z</i>	<i>t</i>
Forget>Remember						
Supramarginal gyrus	R	128210	60	-37	29	8.91
Precentral gyrus	R	115	36	-16	41	4.53
Remember>Forget						
Inferior Frontal gyrus	L	369	-51	17	32	6.75
Inferior Parietal gyrus	L	314	-33	-43	41	6.21
Angular gyrus	R	174	33	-52	44	4.63
Help>Remember						
Paracentral gyrus	R	6094	15	-31	50	7.31
Middle frontal gyrus	R	519	30	50	35	5.73
Inferior frontal gyrus	R	157	48	29	8	5.1
Planum Polare	L	504	-42	-22	-1	6.18
Remember>Help						
Inferior parietal gyrus	L	241	-42	-43	44	6.01
Precentral gyrus	L	330	-45	5	35	5.98
Caudate	L	103	-15	14	8	5.19
Supramarginal gyrus	R	151	45	-31	44	5.91
Forget>Help						
Supramarginal gyrus	L	72	-54	-37	29	5.37
Fusiform gyrus	L	71	-30	-58	-7	5.21

Note. Peak locations of significant differences between the three cue conditions during displaying cues ($p < .05$, FEWR corrected for multiple comparisons at the cluster level with an underlying uncorrected threshold of $p < .001$). No significant differences were observed in the H>F contrast.

5.4 Discussion

We investigated whether brain activity triggered when people receive an offload (or 'help') cue, indicating the presence of a reminder, is similar to when people receive a forget or remember cue. The present study using fMRI recordings demonstrates that help cues induce brain activation that essentially overlap with forget cues, where the regions triggered by a help cue are roughly the reductions in the regions triggered by a forget cue. By contrast, the activation elicited by forget and help cues is prominently different from remember cues. Thus, these results suggest offloading-triggered brain activation is more akin to directed forgetting than remembering information; however, more widespread brain regions are involved in cognitive processes following a forget instruction relative to an offload instruction..

The similarities and differences between directed forgetting and offloading reflect that cognitive processes engaging in offloading in the early phase may be similar to directed forgetting, but with a lower intensity. In particular, the left fusiform gyrus and lingual gyrus, identified in the following analyses as having significant discrepancies between directed forgetting and offloading (i.e. Forget>Help), have been proposed to be crucial structures for the perception of face and scene items (e.g., Kanwisher et al., 1997; Palejwala et al., 2021; Wang et al., 2019). This result of the activation of visual perception regions contributing to the weakening of memory fits with a previous study (Wang et al., 2019) that associated moderate activation of visual sensory regions with the underlying mechanism of directed forgetting. Another region, the left supramarginal gyrus, is also found in our analysis to be involved in cognitive processes underlying the discrepancies between offloading and directed forgetting, and there are at least two possible explanations

for this. First, given that supramarginal gyrus has been proposed to play an important role in storing phonological loop in verbal working memory tasks (e.g., Deschamps et al., 2014; Emach et al., 2019), the results can be explained by the fact that people used phonological loops as a strategy to aid in answering the position test. For example, people could remember a list by subvocally rehearsing the main features of the items sequentially like “short-hair girl, big-eye man, long-hair girl”. Therefore, the higher activation of left supramarginal gyrus in directed forgetting relative to offloading can be associated with phonological memory reduction. Second, supramarginal gyrus has been suggested to be part of the ventral attention system where attention is guided spontaneously to salient relevant stimuli (Stevens et al. 2005; Corbetta et al. 2008). This is evidenced by the consistent findings that attention from temporally and spatially unanticipated stimuli triggers activity in the supramarginal gyrus and temporoparietal junction (e.g., Braver et al., 2001; Indovina & Macaluso, 2006; Stevens et al., 2005; Vossel et al., 2006). It is likely that the attention shift is more pronounced in directed forgetting compared to offloading. The two possibilities are not exclusive mutually, and further analyses on supramarginal gyrus may provide more evidence for the interpretation.

The current study showed that a network of brain regions with higher BOLD signal to a forget cue compared to a remember cue were widespread across bilateral frontal, parietal, temporal and occipital areas. In contrast, the regions implicated by remember instructions appears noticeably less extensive. This suggests that when people intend to forget information rather than retain memories, there are more diverse and broader cognitive processes involved. Specifically, right supramarginal gyrus, in which a peak of activation were identified, has been suggested to engage in processes allowing attention to automatically shift to salient stimuli (Stevens et al.

2005; Corbetta et al. 2008). On the other hand, bilateral inferior parietal gyrus regions show a preferential pattern for a remember cue than a forget cue, playing an important role in the dorsal attention system that volitionally allocates attention. Moreover, given the left inferior frontal gyrus, middle frontal and premotor gyrus were also identified as a region involved in remembering more than directed forgetting, the results are consistent with previous studies revealing prefrontal cortex and inferior parietal co-activation are associated with guiding attention internally to memorial representations (e.g., Kizilirmak et al., 2015; Nee & Jonides, 2009). Thus, remembering information might shift attention to internal memory representations, whereas intentional forgetting could allow attention to be spontaneously directed to external stimuli.

The majority of previous directed forgetting literature showed evidence supporting that frontal-hippocampal inhibitory connectivity plays a crucial role. There are at least three relevant differences between the present experimental methods and previous ones that may provide a more complete picture of the mechanisms involved in directed forgetting. First, we evaluated brain activation during the moment immediately after an instruction was given, whereas most previous fMRI studies have investigated the neural correlates of directed forgetting during List 2 encoding or retrieval phase. It is likely that more cognitive processes engage at the time when people are instructed to forget than during encoding or retrieval phases. For example, from the perspective of context-change accounting for directed forgetting, the cognitive processes involved in the context change may be set up at that instruction phase, while during List 2 encoding and retrieval phases, the outcome of the context change is simply maintained. Second, the methodological and procedural differences between the current study and earlier research may be

related to the absence of finding the inhibitory connectivity. Prior saving-enhanced memory and LMDF studies usually used words as stimuli. In the saving-enhanced memory studies, the items of a list were displayed simultaneously on a screen (Runge et al., 2019; Storm & Stone, 2015) or displayed sequentially one below the other on a single screen (Runge et al., 2021) whereas in LMDF research, the items were presented sequentially on different displays. By contrast, the current study uses images of faces and scenes as stimuli by presenting each item of a list sequentially on different displays. The differences in stimuli and presentation methods can be related to the different neural mechanisms. Finally, using a position test rather than a recall test might contribute to the differences in neural correlates, as the memory processes and possible strategies might be influenced, as exemplified in the above paragraph (i.e. phonological loop). Further analyses specific to frontal and hippocampal regions, such as regions of interest analyses, might help clarify the contradicting results.

5.4.1 Saving-enhanced Memory Effect and Categorical Incompatibility

The behavioural data of the position test revealed the following findings: a) memory performance improvement in the position test of the help trials could demonstrate that provided reminders work to support the unaided memory b) No significant benefit of directed forgetting on the memory of the second list was found when collapsing over compatible and incompatible trials c) It was found that saving the first list led to memory enhancement of the second list, but only on compatible trials d) Memory performance of the second list after remembering the first list was better on incompatible trials than compatible trials. Together, these findings

emphasise that directed forgetting, saving and remembering a previously studied list would have diverse effects on the mnemonic processes of a new list. Moreover, the category compatibility of the two lists could interact with the effect of the three memory-related cognitive processes.

The significant effect of category compatibility on the memory performance of the second list on remember trials could be explained by the mixed-category benefits of working memory. The capacity limits of working memory can be expanded by the utilisation of different category stimuli (e.g., Avital-Cohen & Gronau, 2021; Mruczek et al., 2019). The expanded cognitive resources, namely, the capacity of working memory, could contribute to better memory performance on incompatible trials. On the other hand, categorical incompatibility might facilitate retrieval efficiency through avoiding interference. Since in incompatible trials, the category of the tested item was associated with only one list, the category could serve as context and improve retrieval.

The results that offloading the first list in compatible trials enhanced the memory of the second list fits the previous saving-enhanced memory effect research. However, the saving-enhanced memory effect did not hold on incompatible trials. Instead, people remembered the second list better after remembering the first list than saving it. It is likely that when the two lists are categorically incompatible, the underlying mechanisms of saving-enhanced memory effect such as reset of encoding (i.e. cognitive resource release) or interference reduction (i.e. context inhibition of List 1) would not work (Kanwisher et al., 1997; Palejwala et al., 2021; Pastötter et al., 2012; Runge et al., 2020; Wang et al., 2019). The categorical incompatibility might not only have expanded the capacity of cognitive resource (e.g., working memory)

during encoding the second list, but also avoided the interference in the retrieval. However, a study of Runge et al. (2019) where participants required to solve arithmetic problems after remembering a list found they solved more problems when they offloaded the list. This demonstrates that the saving-enhanced performance effect might be maintained when the cognitive processes involved in the subsequent task were across different cognitive domain. Moreover, given that the memory performance of remember incompatible trials was not the same as help incompatible trials, other mechanisms should exist to bring about the difference. Additionally, the performance differences between forget and remember incompatible trials were not significant, although forget cues and help cues triggered brain activations with resemblances. Further MRI data analyses of the current study or future studies aimed at assessing the interaction between the mixed category effect and the saving-enhanced memory effect may be required to elucidate the possible mechanisms.

Neither of the comparisons between different cues or categorical compatibility revealed any significant differences in the memory performance of the post-MRI memory test, which is a recognition test on all the items displayed in the LMDFS task. While the LMDF paradigm usually includes a final recognition test showing significant performance differences between to-be-forgotten items and to-be-remembered items, previous studies showed List 1 forgetting is usually absent on recognition tests (for reviews, see Anderson & Hanslmayr, 2014; Anderson & Hulbert, 2021; Basden et al., 1993). A LMDF study (Abel et al., 2021) indicates that List 1 forgetting is sustained if participants received a recall test after 20 minutes, whereas the effect of directed forgetting was not found on recognition tests performed either after 30 seconds or after 20 minutes. Therefore, the fact that our results showed no effect of directed

forgetting or saving on memory performance can be explained by test type rather than the effect being short-term. On the other hand, the results might be attributable to the fact that people did not retain the memory of items after a position test due to unawareness of the final memory test; in other words, the forgetting process occurs for all items.

5.4.2 Limitation and Future Research

Given that the cognitive processes involved in offloading may be more similar to directed forgetting, a question remains as to whether the change in activation of the first list content following a help cue would be similar to a forget cue. In order to investigate the question, we can use pattern classifiers in subsequent fMRI analyses to trace the course of brain activation elicited by the contents of the first list. In terms of the neural correlates associated with the benefit effect of directed forgetting and offloading, we plan to compare the brain activation between different cues during encoding of the second list. Critically, as the behavioural data revealed that the saving-enhanced memory effect did not hold on incompatible trials, categorical compatibility should be an important factor in these next MRI analyses so that we can better understand the mechanisms of directed forgetting and offloading.

A potential limitation of the present paradigm comes from the design of giving answers in a straightforward manner on the help trials. The reason for the design is to avoid additional interference, which is quite critical for fMRI studies. However, it is not how cognitive offloading often helps in daily lives. According to the definition of cognitive offloading, people need to perform physical actions to create external tools. Moreover, the design might make the cognitive processes occurring in the help trials more similar to directed forgetting. A future research direction could be a

methodological design that allows the comparison of brain activities triggered by cognitive offloading with and without the incorporation of physical action, which will shed more light on underlying neural correlates of cognitive offloading.

The strength of the current study lies in identifying similarities between lingering activation triggered by offloading cues and either forget cues or remember cues. We found a potentially intuitive result that offloading triggers brain activation more akin to directed forgetting than remembering. Despite the similarities, there are regions implicated in directed forgetting rather than offloading. On the other hand, we found the categorical compatibility of information affects the saving-enhanced memory effect, which highlights the influence of mixed category on the underlying mechanism of cognitive offloading.

6

General Discussion

In this final chapter, I summarise the findings and incorporate them into more general issues in cognitive neuroscience that are associated with the main aims of the studies.

6.1 Overview of the Findings

The empirical work presented in this thesis was developed to address two research questions related to intention offloading:

- 1) How is intention offloading influenced by different factors?
- 2) What are the neural correlates of offloading information?

The first question was addressed with three behavioural experiments.

Experiment 1 evaluated the effect of metacognitive intervention by manipulating a between-subject variable of whether participants received metacognitive advice (i.e. the optimal option). Participants performed an optimal reminder task, where they were financially incentivised to choose to use reminders in an optimal fashion. The results showed that participants were biased towards using external reminders, and the bias was eliminated if participants received metacognitive advice. Experiment 2 used the same task, the optimal reminder task, to investigate the effect of aging on cognitive offloading. Older people used numerically more reminders to compensate for impaired memory abilities compared with younger people. However, based on a comparison between participants' choices and the optimal options determined by weighing costs and benefits of the offloading strategy, older people did not exhibit

pro-reminder bias (i.e. bias toward offloading). The reduction in bias of older people could be related to the finding that older people were particularly underconfident in using external reminders while younger people were particularly underconfident in relying on unaided memory. Experiment 3 investigated how people chose cognitive offloading strategies in a time-based rather than an event-based PM task. Consistent with event-based PM studies, results showed people adjusted offloading intentions according to metacognitive judgment. Factors related to the passage of time, delay length and time-monitoring influenced the extent to which people chose offloading intentions.

To answer the second question, our experiment employed fMRI to compare brain activities between cases where participants were cued to forget, remember or offload information. The study found that the neural correlates triggered by offloading is more similar to directed forgetting than remembering. However, there were some regions of brain activation associated with directed forgetting rather than offloading, suggesting that differences in cognitive processes may lie behind these two phenomena.

6.2 Cognitive Offloading in PM Tasks

To the best of my knowledge, it is novel that experiment 1 and 2 used the optimal reminder task to evaluate not only the extent to which people offload intentions, but also how optimally people offload intentions. Results showed people were inclined to use more reminders than they actually needed, and the pro-reminder bias was not found in older adults. This does not mean that it is always the case in everyday life that people overuse reminders and older adults use reminders

more optimally. To be clear, the results are in line with a plurality of prior offloading research (Dunn & Risko, 2016; Gilbert, 2015a, 2015b; Gilbert et al., 2020) suggesting that people modify offloading intentions based on various cognitive processes such as metacognition or inner preference. Given that most previous past studies have investigated intention offloading in event-based PM tasks, Experiment 3 evaluated offloading intentions in a time-based PM task and found that cognitive processes such as metacognitive judgment or avoiding cognitive effort affecting offloading were common to both event-based and time-based tasks.

6.3 Factors influencing cognitive offloading

The thesis evaluated the effects of metacognitive advice, aging processes, and PM task type on intention offloading in PM tasks. Given the processes of offloading intentions involve a variety of cognitive processes which may be targeted by intrinsic (i.e. aging) and extrinsic (i.e. PM task type, metacognitive advice) factors, the effect of these factors will be discussed in a twofold manner: a) the modification of cognitive processes influences strategic offloading decisions and b) various factors affect these cognitive processes, which in turn influence offloading decisions. Having proposed that metacognitive processes and aversion to cognitive effort are the essential cognitive processes integrated in the decision about intention offloading, we will discuss the consequences of the intrinsic and extrinsic factors based on metacognitive judgment and the aversion to cognitive effort acting as mediators.

6.3.1 Metacognition

Studies have revealed that metacognitive processes play a crucial role in

offloading intentions regardless of age group or PM task type, and this allows metacognitive advice to help optimise the use of offloading strategies. Experiment 1 demonstrated the extent to which people were biased to predict poorer unaided memory ability, leading to the extent to which they set reminders more than they actually needed. Participants receiving metacognitive advice exhibited no pro-reminder bias, implying that metacognitive advice guides people to optimise offloading decisions. Furthermore, Experiment 3 showed that the effects of metacognitive processes were found when comparing strategic offloading decisions across subjects and across delays, finding not only did individuals who have less confidence in unaided memory set more reminders, but also finding that individuals with a greater drop in confidence at longer delays were more likely to set reminders when encountering longer delays. Additionally, In Experiment 2, despite having no direct evidence (i.e. significant correlation between the measures of metacognitive judgment and reminder usage), a reduction in reminder bias in older adults could be explained by the resulting upward shift in confidence about unaided memory, relative to actual performance, as well as underestimating the benefit effects of reminders on memory performance. This highlights that, in addition to metacognitive judgment about unaided memory, other related metacognitive processes such as metacognitive judgment about reminder benefit may impact offloading intentions.

Having shown the effect of metacognitive beliefs on offloading intentions, the studies also demonstrate how different factors affect metacognitive processes. Experiment 2 found that older adults had upward shifts in confidence about unaided memory, relative to actual performance, is consistent with other studies demonstrating increased overconfidence in older participants (Bruce et al., 1982; Cauvin et al., 2019; Connor et al., 1997; Scarampi & Gilbert, 2021; Soderstrom et al.,

2012). This is possibly explained by failing to update metacognitive beliefs (cf. Knight et al, 2005; Souchay, 2007), and may be related to age-related decline in cognitive abilities. While the intrinsic factor, aging, can bring general changes in metacognitive processes, manipulating the extrinsic factors, metacognitive advice and PM task type, shows how metacognitive processes are adapted in response to various external cues (i.e. metacognitive advice, delay length in the time-based PM task). Even though no direct evidence (i.e. confidence measures after giving advice) was provided in Experiment 1 supporting that people's metacognitive judgment was modified based on the advice, a few other intention offloading studies (Engeler & Gilbert, 2020; Sachdeva & Gilbert 2020; Gilbert et al, 2020 Experiment 3) have demonstrated that metacognitive judgment is influenced by positive and negative feedbacks, altering strategic offloading decisions. Moreover, Experiment 3 found that the modification of metacognitive judgments about unaided memory ability across different delays was not predicted by their objective, unaided accuracy across different delays. However, the objective, unaided accuracy averaged across delays predicted confidence averaged across delays. It was likely that the modification of metacognitive judgments across different delays was influenced by extra factors other than actual performance, such as subjective beliefs about the extent to which delay length could lead to the changes in memory performance. Thus, our results suggest that metacognitive processes can be influenced by external cues that are part of the extrinsic environment, which plays an important role in the processes of intention offloading.

6.3.2 Cognitive load

The thesis provides some evidence supporting that aversion to cognitive effort is

another integral process associated with offloading intentions. First, Experiment 2 found that although older adults particularly underestimated the benefit of offloading, they did not have a bias towards unaided memory (i.e. using fewer reminders than they actually needed), which suggests that one or more additional factors contributed to strategic offloading decisions. Sachdeva and Gilbert (2020) proposed one of the factors could be aversion to cognitive effort. They found that using financial incentive to encourage participants' offloading intentions optimally reduced the pro-reminder bias, which can be explained by the likelihood that the financial incentive increased participants' willingness to spend cognitive effort on the task. Second, Experiment 4 revealed that reduced clock revealability (i.e. whether the clock is displayed constantly) contributed to unique variance of reminder usage when metacognitive judgment was controlled, which might be due to increased cognitive demand brought about by clock revealability. This implies that greater cognitive demand of PM tasks leads to an increase in offloading intentions in order to avoid cognitive effort.

Despite the lack of objective assessment of changes in cognitive demand, extrinsic factors clearly impact the amount of cognitive effort required to complete PM tasks. In Experiment 4, people were proposed to expend more cognitive effort when the delay was longer or the clock was hidden in a time-based PM task. In Experiment 1, providing metacognitive advice was hypothesised to reduce cognitive effort required for metacognitive processes. This reduction in cognitive effort could lead advised participants to follow advice, and therefore showed no bias towards offloading. On the other hand, intrinsic factors such as aging might adjust the extent to which people avoid cognitive effort (Ennis et al., 2013), or alter cognitive capability which, in turn, changes the amount of cognitive effort required for tasks or setting

reminders (e.g., memory ability decline or individuals with lower memory ability; Ball et al., 2021; Scarampi & Gilbert, 2021). Experiment 2 suggests that the absence of pro-reminder bias in older adults can be explained by the likelihood that although setting reminder reduces cognitive demand, older people may consider the additional reminder-setting behaviour to be more effortful due to the requirement of switching away from the ongoing task to setting reminders (see Chapter 3 for discussion).

6.4 Neural Correlates of Cognitive Offloading

The results of Experiment 4 shows that the brain activation of offloading is similar to directed forgetting. These results align to previous studies where offloading information led to the reduction of memory for that information (Eskritt & Ma, 2013; Grinschgl et al., 2021; Henkel, 2014; Kang, 2022; Lurie & Westerman, 2021; Sparrow et al., 2011; Storm & Stone, 2015; Tamir et al., 2018). Sparrow et al. (2011) defined people's tendency to forget the information they believe will be accessible later (e.g., from web search engines) as "Google effects on memory". Ward (2013) suggested that people have greater confidence in their internal memory ability when they believe the information would be accessible later through search engine. Given the crucial role metacognitive processes play in strategic offloading decisions, it is likely that having access to a search engine could lead people to overestimate their cognitive ability and offload information less than they should.

Previous directed forgetting literature (for a review see Anderson & Hanslmayr, 2014; Anderson & Hulbert, 2021) has defined directed forgetting as active modification in memory processes by gauging whether any temporarily unnecessary

information should be discarded or suppressed through specific mechanisms in order to avoid the overload of cognitive abilities. The brain regions triggered by directed forgetting or offloading compared to a remembering were more widespread, potentially suggesting that forgetting or offloading involve more diverse and broader cognitive processes than retaining memory. Specifically, while the regions engaged in shifting attention to salient environment stimuli were more activated after forgetting and offloading cues, remembering previous displayed stimuli showed activation related to the processes that maintain attention to existing memory representatives. This may indicate that transforming a cognitive state into a better preparedness to remember new information involves a wide range of cognitive processes such as attention processes.

6.5 Beneficial Effects of Cognitive Offloading

Past studies have demonstrated that the benefits of cognitive offloading include not only improved memory performance of saved information, but also reduced cognitive demand for internal memory (Herrmann et al., 1999), which could in turn improve the performance of subsequent tasks (e.g., Runge et al., 2019; Storm & Stone, 2015). The current thesis supports this view with three main findings. First, the thesis including behavioural data of prospective and retrospective memory studies showed that people performed better when using offloading strategies than relying on memory alone. Second, increasing offloading in a PM task may not only result in better performance on the PM task but also the ongoing task. Experiment 3 showed that the hidden-clock group (i.e. the clock was not constantly displayed) allowed to set reminders, got a greater accuracy on both the PM task and ongoing task. Finally, Experiment 4 showed categorical compatibility (i.e. saved and unsaved

information belonging to the same or different category) influences the saving-enhanced memory effect. This led directly to the question of whether the improvement of subsequent task performance could hold across category or across domain.

6.6 Predicted and Unpredicted Results

In Experiment 1, our hypothesis was that if the bias in participants' choices was solely driven by metacognitive judgments, then providing metacognitive advice about performance could reduce the bias. The results demonstrated that the reminder bias was eliminated in the group that received advice. Furthermore, the correlation between metacognitive bias and reminder bias was significant only in unadvised groups.

In Experiment 2, we hypothesised that older individuals would exhibit lower confidence in their unaided memory ability and display more reminder bias compared to younger individuals. The results revealed that older participants did set more reminders numerically, but their reminder bias was reduced when their impaired memory ability was taken into account. In terms of confidence, older individuals were less confident in their memory performance, but their confidence was upwardly adjusted relative to their actual performance. Additionally, we unexpectedly found that older participants underestimated the beneficial effects of reminders.

In Experiment 3, the predicted result was that people relied on more reminders when the delay was longer, and when the clock was not persistently visible. Additionally, we predicted that the impact of delay length on offloading behaviour would be modulated by metacognitive judgment. We did not have a specific

hypothesis regarding the influence of delay length and clock visibility on unaided memory performance. Nevertheless, the results revealed that participants demonstrated worse unaided memory performance when the delay was longer. Another unexpected finding was that participants' overall metacognitive judgment was not significantly associated with their general tendency to offload. In Experiment 4, our hypothesis was that offloading would be different from directed forgetting and aimed to identify the neural correlates which could differentiate between the two processes. The results revealed that brain activation triggered by offloading was more akin to directed forgetting than to remembering. This suggested a similarity in cognitive processes between directed forgetting and cognitive offloading, although they still differed from each other.

6.7 Limitation and Suggestions for Future Research

In this thesis, the optimal reminder task was used to evaluate whether people offloaded intentions optimally. The optimal reminder task is fairly difficult, where participants usually have an accuracy rate of about fifty percent using unaided memory. Moreover, the financial incentive encourages people offloading strategies optimally, which could enhance the effect of metacognitive judgment. However, most PM tasks in daily life are not so difficult and are not related to monetary enticement. Therefore, more consideration should be given to other factors such as avoiding cognitive effort or motivation when applying the results to daily use of offloading strategies in PM tasks.

This thesis aims to explore whether certain factors (i.e. metacognitive advice,

aging and PM task type) influence intention offloading and whether metacognitive judgments about memory performance are crucial for strategically offloading decisions. A plausible path for pursuing the answers to these questions could be to investigate the effects of other relevant cognitive processes such as avoiding cognitive effort, strategy perseveration or motivation through designing objective assessment for or manipulating these processes. Moreover, more evidence is needed to establish a theoretical account for the underlying mechanisms of strategic offloading decisions, which can help understand how various cognitive processes integrate with each other.

Given that offloading information to external devices may decrease the memory performance for the offloaded information, and improve memory performance for completing subsequent tasks, it will be necessary to find out how to avoid the influence from detrimental memory consequences and amplify beneficial memory consequences. For example, using GPS may make it difficult for people to identify the routes they drove before, but allow people to find their way easier and focus on driving. It would be detrimental if people who consistently utilize GPS believe they can still recognise the routes without the help of GPS, or if people choose to rely GPS but it breaks down on the road. Further research could evaluate whether metacognitive intervention could rectify the erroneous metacognitive judgments triggered by offloading or distinguish tasks that should not be offloaded.

Because of today's close integration of technology into everyday life, these insights in offloading decisions, and the cognitive consequences of using offloading strategies, can help explore better advancement in this digital environment. First, digital applications could provide metacognitive intervention through devices that

record and feedback the outcomes of previous retrospective or prospective memory tasks. Also, our knowledge about offloading strategies could be incorporated into new applications. For example, devices can provide reinforcements to prevent older people from not setting enough reminders to compensate memory decline. Furthermore, devices need to be designed with the goal of making it easier for older adults to learn how to set reminders.

6.8 Conclusions

The current thesis derives the following insights from the behavioural and fMRI experiments. The results suggest that:

1) People do not always offload intentions optimally, and they adjust intention offloading based on the modification of metacognitive judgements about unaided memory, and cognitive load in both event-based, and time-based PM tasks.

2) Intrinsic (i.e. aging processes) and extrinsic (i.e. metacognitive intervention) factors impact cognitive offloading through affecting underlying cognitive processes such as metacognition.

3) Brain activation triggered by offloading is more similar to directed forgetting relative to remembering, suggesting the resemblance of cognitive processes between cognitive offloading and directed forgetting.

References

- Aarts, H., Ruys, K. I., Veling, H., Renes, R. A., de Groot, J. H. B., van Nunen, A. M., & Geertjes, S. (2010). The art of anger: Reward context turns avoidance responses to anger-related objects into approach. *Psychological Science*, *21*(10), 1406–1410. <https://doi.org/10.1177/0956797610384152>
- Abel, M., Kuchler, B., Meier, E., & Bäuml, K.-H. T. (2021). List-method directed forgetting: Do critical findings generalize from short to long retention intervals? *Memory & Cognition*, *49*(8), 1677–1689. <https://doi.org/10.3758/s13421-021-01192-z>
- 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*(6), 1444–1453. <https://doi.org/10.1037/a0020718>
- Anderson, M. C., Bunce, J. G., & Barbas, H. (2016). Prefrontal–hippocampal pathways underlying inhibitory control over memory. *Neurobiology of Learning and Memory*, *134*(Pt A), 145–161. <https://doi.org/10.1016/j.nlm.2015.11.008>
- Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature (London)*, *410*(6826), 366–369. <https://doi.org/10.1038/35066572>
- Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences*, *18*(6), 279–292. <https://doi.org/10.1016/j.tics.2014.03.002>
- Anderson, M. C., & Hulbert, J. C. (2021). Active forgetting: Adaptation of memory by prefrontal control. *Annual Review of Psychology*, *72*(1), 1–36. <https://doi.org/10.1146/annurev-psych-072720-094140>
- Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., Glover, G. H., & Gabrieli, J. D. E. (2004). Neural systems underlying the suppression of unwanted memories. *Science (American Association for the Advancement of Science)*, *303*(5655), 232–235. <https://doi.org/10.1126/science.1089504>
- Apps, M. A. J., Grima, L. L., Manohar, S., & Husain, M. (2015). The role of cognitive effort in subjective reward devaluation and risky decision-making. *Scientific Reports*, *5*, 1–11. <https://doi.org/10.1038/srep16880>
- Arango-Muñoz, S. (2013). Scaffolded memory and metacognitive feelings. *Review of Philosophy and Psychology*, *4*(1), 135–152. <https://doi.org/10.1007/s13164-012-0124-1>
- Avital-Cohen, R., & Gronau, N. (2021). The asymmetric mixed-category advantage in

- visual working memory: An attentional, not perceptual (face-specific) account. *Journal of Experimental Psychology. Human Perception and Performance*, 47(6), 852–868. <https://doi.org/10.1037/xhp0000921>
- Azzopardi, B., Auffray, C., & Kermarrec, C. (2017). Effet paradoxal du vieillissement sur des tâches de mémoire prospective *time-based* en situation de laboratoire et en milieu naturel: Rôle des fonctions exécutives. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement*, 36(1), 30–40. <https://doi.org/10.1017/S0714980816000738>
- Baird, B., Smallwood, J., Gorgolewski, K. J., & Margulies, D. S. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *Journal of Neuroscience*, 33(42), 16657–16665. <https://doi.org/10.1523/JNEUROSCI.0786-13.2013>
- Ball, H., Peper, P., Alakbarova, D., Brewer, G., & Gilbert, S. J. (2021). Individual differences in working memory capacity predict benefits to memory from intention offloading. *Memory*, 1–15. <https://doi.org/10.1080/09658211.2021.1991380>
- Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: a comparison of methods. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 19(3), 603–616. <https://doi.org/10.1037/0278-7393.19.3.603>
- Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, 16(6), 351–355. <https://doi.org/10.1111/j.1467-8721.2007.00534.x>
- Bäuml, K.-H., Hanslmayr, S., Pastötter, B., & Klimesch, W. (2008). Oscillatory correlates of intentional updating in episodic memory. *NeuroImage (Orlando, Fla.)*, 41(2), 596–604. <https://doi.org/10.1016/j.neuroimage.2008.02.053>
- Benoit, R. G., Hulbert, J. C., Huddleston, E., & Anderson, M. C. (2015). Adaptive top-down suppression of hippocampal activity and the purging of intrusive memories from consciousness. *Journal of Cognitive Neuroscience*, 27(1), 96–111. https://doi.org/10.1162/jocn_a_00696
- Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of abbreviated nine-item forms of the Raven's Standard Progressive Matrices Test. *Assessment*, 19(3), 354–369. <https://doi.org/10.1177/1073191112446655>
- Birchill, W. R., & Schaie, K. W. (1975). The effect of differential reinforcement of cautiousness in intellectual performance among the elderly. *Journal of Gerontology*, 30(5), 578–583. <https://doi.org/10.1093/geronj/30.5.578>
- 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(6), 569–577. <https://doi.org/10.1007/BF03324886>
- Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In *Varieties of memory and consciousness: Essays in honour of Endel Tulving*. (pp. 309–330). Lawrence Erlbaum Associates, Inc.
- Boldt, A., & Gilbert, S. J. (2019b). Confidence guides spontaneous cognitive offloading. *Cognitive Research: Principles and Implications*, 4(1). <https://doi.org/10.1186/s41235-019-0195-y>
- Brandimonte M. A., Ferrante D., Feresin C., & Delbello R. (2001). Dissociating prospective memory from vigilance processes? *Psicológica (Valencia)*, 22(1), 97–113.
- Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of frequency, inhibition and errors. *Cerebral Cortex (New York, N.Y. 1991)*, 11(9), 825–836. <https://doi.org/10.1093/cercor/11.9.825>
- Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, 10(1), 12–21. <https://doi.org/10.1080/17470215808416249>
- Bruce, P. R., Coyne, A. C., & Botwinick, J. (1982). Adult age differences in metamemory. *Journal of Gerontology*, 37(3), 354–357. <https://doi.org/10.1093/geronj/37.3.354>
- Brügger, A., Richter, K.-F., & Fabrikant, S. I. (2019). How does navigation system behavior influence human behavior? *Cognitive Research: Principles and Implications*, 4(1), 5–5. <https://doi.org/10.1186/s41235-019-0156-5>
- Burgess, P. W. (2000). Strategy application disorder: The role of the frontal lobes in human multitasking. *Psychological Research*, 63(3–4), 279–288. <https://doi.org/10.1007/s004269900006>
- Castel, A. D., McGillivray, S., & Middlebrooks, C. D. (2016). *Monitoring memory in old age: Impaired, spared, and aware*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199336746.013.3>
- Castel, A. D., Murayama, K., Friedman, M. C., McGillivray, S., & Link, I. (2013). Selecting valuable information to remember: Age-related differences and similarities in self-regulated learning. *Psychology and Aging*, 28, 232–242. <https://doi.org/10.1037/a0030678>
- Castel, A. D., Rhodes, M. G., & Friedman, M. C. (2013). Predicting memory benefits in the production effect: The use and misuse of self-generated distinctive cues when making judgments of learning. *Memory & Cognition*, 41(1), 28–35. <https://doi.org/10.3758/s13421-012-0249-6>

- Causey, K. B., & Bjorklund, D. F. (2014). Prospective memory in preschool children: Influences of agency, incentive, and underlying cognitive mechanisms: The development of prospective memory in childhood. *Journal of Experimental Child Psychology*, *127*, 36–51.
- Cauvin, S., Moulin, C. J. A., Souchay, C., Kliegel, M., & Schnitzspahn, K. M. (2019). Prospective memory predictions in aging: Increased overconfidence in older adults. *Experimental Aging Research*, *45*(5), 436–459.
<https://doi.org/10.1080/0361073X.2019.1664471>
- Chen, K., & Chan, A. H. S. (2014). Gerontechnology acceptance by elderly Hong Kong Chinese: A senior technology acceptance model (STAM). *Ergonomics*, *57*(5), 635–652. <https://doi.org/10.1080/00140139.2014.895855>
- Cherkaoui, M., & Gilbert, S. J. (2017). Strategic use of reminders in an ‘intention offloading’ task: Do individuals with autism spectrum conditions compensate for memory difficulties? *Neuropsychologia*, *97*(January), 140–151.
<https://doi.org/10.1016/j.neuropsychologia.2017.02.008>
- Cicerone, K. D., Langenbahn, D. M., Braden, C., Malec, J. F., Kalmar, K., Fraas, M., Felicetti, T., Laatsch, L., Harley, J. P., Bergquist, T., Azulay, J., Cantor, J., & Ashman, T. (2011). Evidence-based cognitive rehabilitation: Updated review of the literature from 2003 through 2008. *Archives of Physical Medicine and Rehabilitation*, *92*(4), 519–530. <https://doi.org/10.1016/j.apmr.2010.11.015>
- Clark, A. (1996). *Being There: Putting Brain, Body, and World Together Again*. The MIT Press. <https://doi.org/10.7551/mitpress/1552.001.0001>
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7–19.
<https://doi.org/10.1111/1467-8284.00096>
- Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*, *12*(1), 50–71.
<https://doi.org/10.1037/0882-7974.12.1.50>
- Conte, A. M., & McBride, D. M. (2018). Comparing time-based and event-based prospective memory over short delays. *Memory*, *26*(7), 936–945.
<https://doi.org/10.1080/09658211.2018.1432662>
- Cook, G. I., Marsh, R. L., Clark-Foos, A., & Thadeus Meeks, J. (2007). Learning is impaired by activated intentions. *Psychonomic Bulletin and Review*, *14*(1), 101–106. <https://doi.org/10.3758/BF03194035>
- Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human Memory and Cognitive Capabilities: Mechanisms and Performances* (pp. 409–422). Elsevier.
- Craik, F. I. M., & Kerr, S. A. (1996). Commentary: Prospective memory, aging, and lapses of intention. In G. O. Einstein & M. A. Mark (Eds.), *Prospective*

- memory: Theory and applications* (pp. 227–237). Lawrence Erlbaum.
- de Gardelle, V., & Mamassian, P. (2014). Does confidence use a common currency across two visual tasks? *Psychological Science*, *25*(6), 1286–1288.
<https://doi.org/10.1177/0956797614528956>
- Depue, B. E., Curran, T., & Banich, M. T. (2007). Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science (American Association for the Advancement of Science)*, *317*(5835), 215–219.
<https://doi.org/10.1126/science.1139560>
- Detre, G. J., Natarajan, A., Gershman, S. J., & Norman, K. A. (2013). Moderate levels of activation lead to forgetting in the think/no-think paradigm. *Neuropsychologia*, *51*(12), 2371–2388.
<https://doi.org/10.1016/j.neuropsychologia.2013.02.017>
- Devolder, P. A., Brigham, M. C., & Pressley, M. (1990). Memory performance awareness in younger and older adults. *Psychology and Aging*, *5*(2), 291–303.
<https://doi.org/10.1037//0882-7974.5.2.291>
- Dobbs, A. R., & Reeves, M. B. (1996). Prospective memory: More than memory. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications*. (pp. 199–225). Lawrence Erlbaum Associates Publishers.
- Dobbs, A. R., & Rule, B. G. (1987). Prospective memory and self-reports of memory abilities in older adults. *Canadian Journal of Psychology*, *41*(2), 209–222.
<https://doi.org/10.1037/h0084152>
- Dreisbach, G., & Fischer, R. (2015). Conflicts as aversive signals for control adaptation. *Current Directions in Psychological Science : A Journal of the American Psychological Society*, *24*(4), 255–260.
- Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. SAGE Publications, Incorporated.
- Dunn, T. L., & Risko, E. F. (2016). Toward a metacognitive account of cognitive offloading. *Cognitive Science*, *40*(5), 1080–1127.
<https://doi.org/10.1111/cogs.12273>
- Dupont, D., Zhu, Q., & Gilbert, S. J. (2022). Value-based routing of delayed intentions into brain-based versus external memory stores. *Journal of Experimental Psychology. General*. <https://doi.org/10.1037/xge0001261>
- Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *16*(4), 717–726.
- 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(4), 996–1007. <https://doi.org/10.1037/0278-7393.21.4.996>
- Einstein, G. O., McDaniel, M. A., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., & Breneiser, J. (2005). Multiple processes in prospective memory retrieval: Factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology: General*, 134(3), 327–342. <https://doi.org/10.1037/0096-3445.134.3.327>
- Einstein, G. O., McDaniel, M. A., Williford, C. L., Pagan, J. L., & Dismukes, R. K. (2003). Forgetting of intentions in demanding situations is rapid. *Journal of Experimental Psychology: Applied*, 9(3), 147–162. <https://doi.org/10.1037/1076-898X.9.3.147>
- Engeler, N. C., & Gilbert, S. J. (2020). The effect of metacognitive training on confidence and strategic reminder setting. *PLOS ONE*, 15(10), e0240858. <https://doi.org/10.1371/journal.pone.0240858>
- Ennis, G. E., Hess, T. M., & Smith, B. T. (2013). The impact of age and motivation on cognitive effort: Implications for cognitive engagement in older adulthood. *Psychology and Aging*, 28(2), 495–504. <https://doi.org/10.1037/a0031255>
- Eskritt, M., & Ma, S. (2013). Intentional forgetting: Note-taking as a naturalistic example. *Memory & Cognition*, 42(2), 237–246. <https://doi.org/10.3758/s13421-013-0362-1>
- Fawcett, J. M., & Taylor, T. L. (2008). Forgetting is effortful: Evidence from reaction time probes in an item-method directed forgetting task. *Memory & Cognition*, 36(6), 1168–1181. <https://doi.org/10.3758/MC.36.6.1168>
- Fawcett, J. M., & Taylor, T. L. (2010). Directed forgetting shares mechanisms with attentional withdrawal but not with stop-signal inhibition. *Memory & Cognition*, 38(6), 797–808. <https://doi.org/10.3758/MC.38.6.797>
- Ferguson, A. M., McLean, D., & Risko, E. F. (2015). Answers at your fingertips: Access to the Internet influences willingness to answer questions. *Consciousness and Cognition*, 37, 91–102. <https://doi.org/10.1016/j.concog.2015.08.008>
- Festini, S. B., & Reuter-Lorenz, P. A. (2017). Rehearsal of to-be-remembered items is unnecessary to perform directed forgetting within working memory: Support for an active control mechanism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(1), 94–108. <https://doi.org/10.1037/xlm0000308>
- Fisher, M., Goddu, M. K., & Keil, F. C. (2015). Searching for explanations: How the internet inflates estimates of internal knowledge. *Journal of Experimental Psychology: General*, 144(3), 674–687. <https://doi.org/10.1037/xge0000070>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of

- cognitive-developmental inquiry. *The American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Fleming, S. M., Dolan, R. J., & Frith, C. D. (2012). Metacognition: Computation, biology and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1280–1286. <https://doi.org/10.1098/rstb.2012.0021>
- Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal contributions to metacognition in perceptual decision making. *Journal of Neuroscience*, 32(18), 6117–6125. <https://doi.org/10.1523/JNEUROSCI.6489-11.2012>
- Gagnepain, P., Henson, R. N., & Anderson, M. C. (2014). Suppressing unwanted memories reduces their unconscious influence via targeted cortical inhibition. *Proceedings of the National Academy of Sciences - PNAS*, 111(13), E1310–E1319. <https://doi.org/10.1073/pnas.1311468111>
- Gagnepain, P., Hulbert, J., & Anderson, M. C. (2017). Parallel regulation of memory and emotion supports the suppression of intrusive memories. *The Journal of Neuroscience*, 37(27), 6423–6441. <https://doi.org/10.1523/JNEUROSCI.2732-16.2017>
- Geiselman, R. E., Bjork, R. A., & Fishman, D. L. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General*, 112(1), 58–72. <https://doi.org/10.1037/0096-3445.112.1.58>
- Gilbert, S. J. (2011). Decoding the content of delayed intentions. *The Journal of Neuroscience*, 31(8), 2888–2894. <https://doi.org/10.1523/JNEUROSCI.5336-10.2011>
- Gilbert, S. J. (2015a). Strategic offloading of delayed intentions into the external environment. *Quarterly Journal of Experimental Psychology*, 68(5), 971–992. <https://doi.org/10.1080/17470218.2014.972963>
- Gilbert, S. J. (2015b). Strategic use of reminders: Influence of both domain-general and task-specific metacognitive confidence, independent of objective memory ability. *Consciousness and Cognition*, 33, 245–260. <https://doi.org/10.1016/j.concog.2015.01.006>
- Gilbert, S. J., Bird, A., Carpenter, J., Fleming, S. M., Sachdeva, C., & Tsai, P.-C. (2020). Optimal use of reminders: Metacognition, effort, and cognitive offloading. *Journal of Experimental Psychology: General*, 149(3), 501–517. <https://doi.org/10.1037/xge0000652>
- Gilbert, S. J., Boldt, A., Sachdeva, C., Scarampi, C., & Tsai, P.-C. (2022). Outsourcing memory to external tools: A review of ‘intention offloading’. *Psychonomic bulletin & review*, 2022 Jul 5. doi: 10.3758/s13423-022-02139-4. Epub ahead

- of print. PMID: 35789477.
- Glisky, E. L. (1996). Prospective memory and the frontal lobes. In *Prospective memory: Theory and applications*. (pp. 249–266). Lawrence Erlbaum Associates Publishers.
- Graf, P., & Uttl, B. (2001). Prospective memory: a new focus for research. *Consciousness and Cognition, 10*(4), 437–450.
<https://doi.org/10.1006/ccog.2001.0504>
- Gray, W. D., Sims, C. R., Fu, W. T., & Schoelles, M. J. (2006). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review, 113*(3), 461–482.
<https://doi.org/10.1037/0033-295X.113.3.461>
- Grinschgl, S., Papenmeier, F., & Meyerhoff, H. S. (2021). Consequences of cognitive offloading: Boosting performance but diminishing memory. *Quarterly Journal of Experimental Psychology (2006), 74*(9), 1477–1496.
<https://doi.org/10.1177/17470218211008060>
- Guajardo, N. R., & Best, D. L. (2000). Do preschoolers remember what to do? Incentive and external cues in prospective memory. *Cognitive Development, 15*(1), 75–97. [https://doi.org/10.1016/S0885-2014\(00\)00016-2](https://doi.org/10.1016/S0885-2014(00)00016-2)
- Haas, M., Mehl, M. R., Ballhausen, N., Zuber, S., Kliegel, M., & Hering, A. (2022). The sounds of memory: Extending the age-prospective memory paradox to everyday behavior and conversations. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 77*(4), 695–703.
<https://doi.org/10.1093/geronb/gbac012>
- Haas, M., Zuber, S., Kliegel, M., & Ballhausen, N. (2020). Prospective memory errors in everyday life: Does instruction matter? *Memory, 28*(2), 196–203.
<https://doi.org/10.1080/09658211.2019.1707227>
- Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin, 136*(4), 495–525. <https://doi.org/10.1037/a0019486>
- Hanslmayr, S., Volberg, G., Wimber, M., Oehler, N., Staudigl, T., Hartmann, T., Raabe, M., Greenlee, M. W., & Bäuml, K.-H. T. (2012). Prefrontally driven downregulation of neural synchrony mediates goal-directed forgetting. *The Journal of Neuroscience, 32*(42), 14742–14751.
<https://doi.org/10.1523/JNEUROSCI.1777-12.2012>
- Harris, J. E. (1984). Remembering to do things: A forgotten topic. In J. Harris & P. Morris (Eds.), *Everyday memory: Actions and absent mindedness* (pp. 71–92). Academic Press.
- Harris, J. E., & Wilkins, A. (1982). Remembering to do things: A theoretical

- framework and illustrative experiment. *Human Learning*, 1.
- Heersmink, R., & Sutton, J. (2020). Cognition and the web: Extended, transactive, or scaffolded? *Erkenntnis*, 85(1), 139–164. <https://doi.org/10.1007/s10670-018-0022-8>
- Henkel, L. A. (2014). Point-and-shoot memories: The influence of taking photos on memory for a museum tour. *Psychological Science*, 25(2), 396–402. <https://doi.org/10.1177/0956797613504438>
- Henry, J. D., MacLeod, M. S., Phillips, L. H., & Crawford, J. R. (2004c). A meta-analytic review of prospective memory and aging. *Psychology and Aging*, 19(1), 27–39. <https://doi.org/10.1037/0882-7974.19.1.27>
- 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. *Quarterly Journal of Experimental Psychology*, 65(7), 1274–1287. <https://doi.org/10.1080/17470218.2011.651091>
- 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: JINS*, 24(6), 640–645. <https://doi.org/10.1017/S1355617718000152>
- Hering, A., Phillips, L. H., & Kliegel, M. (2014). Importance effects on age differences in performance in event-based prospective memory. *Gerontology*, 60(1), 73–78. <https://doi.org/10.1159/000355057>
- Herrmann, D., Brubaker, B., Yoder, C., Sheets, V., & Tio, A. (1999). Devices that remind. In *Handbook of applied cognition*. (pp. 377–407). John Wiley & Sons Ltd.
- Hertzog, C. (2016). Aging and metacognitive control. In J. Dunlosky & S. K. Tauber (Eds.), *The Oxford Handbook of Metamemory*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199336746.013.31>
- Hertzog, C., Dunlosky, J., & Sinclair, S. M. (2010). Episodic feeling-of-knowing resolution derives from the quality of original encoding. *Memory and Cognition*, 38(6), 771–784. <https://doi.org/10.3758/MC.38.6.771>
- Hertzog, C., Lövdén, M., Lindenberger, U., & Schmiedek, F. (2017). Age differences in coupling of intraindividual variability in mnemonic strategies and practice-related associative recall improvements. *Psychology and Aging*, 32(6), 557–571. <https://doi.org/10.1037/pag0000177>
- Hertzog, C., Lustig, E., Pearman, A., & Waris, A. (2019). Behaviors and strategies supporting everyday memory in older adults. *Gerontology*, 65(4), 419–429. <https://doi.org/10.1159/000495910>
- Hertzog, C., Price, J., & Dunlosky, J. (2012). Age differences in the effects of

- experimenter-instructed versus self-generated strategy use. *Experimental Aging Research*, 38(1), 42–62.
<https://doi.org/10.1080/0361073X.2012.637005>
- Hicks, J. L., Franks, B. A., & Spitler, S. N. (2017). Prior task experience and comparable stimulus exposure nullify focal and nonfocal prospective memory retrieval differences. *Quarterly Journal of Experimental Psychology (2006)*, 70(10), 1997–2006. <https://doi.org/10.1080/17470218.2016.1217891>
- Holbrook, J., & Dismukes, K. (2009). Prospective memory in everyday tasks. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 53(10), 590–594.
- Honeywell, J. A., Dickinson, A. M., & Poling, A. (1997). INDIVIDUAL PERFORMANCE AS A FUNCTION OF INDIVIDUAL AND GROUP PAY CONTINGENCIES. *The Psychological Record*, 47(2), 261–274. <https://doi.org/10.1007/BF03395224>
- Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 26, 107–129. [https://doi.org/10.1016/0001-6918\(67\)90011-X](https://doi.org/10.1016/0001-6918(67)90011-X)
- 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(1–2), 84–101. <https://doi.org/10.1080/13825585.2011.629288>
- Indovina, I., & Macaluso, E. (2006). Dissociation of stimulus relevance and saliency factors during shifts of visuospatial attention. *Cerebral Cortex (New York, N.Y. 1991)*, 17(7), 1701–1711. <https://doi.org/10.1093/cercor/bhl081>
- Inzlicht, M., Shenav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly and valued. *Trends in Cognitive Sciences*, 22(4), 337–349.
<https://doi.org/10.1016/j.tics.2018.01.007>
- Johnson, H. M. (1994). Processes of successful intentional forgetting. *Psychological Bulletin*, 116(2), 274–292. <https://doi.org/10.1037/0033-2909.116.2.274>
- Joly-Burra, E., Linden, M. V., & Ghisletta, P. (2018). Intraindividual variability in inhibition and prospective memory in healthy older adults: Insights from response regularity and rapidity. *Journal of Intelligence*, 6(1), 13.
<http://dx.doi.org/10.3390/jintelligence6010013>
- Jones, W. E., Bengel, J. F., & Scullin, M. K. (2021). Preserving prospective memory in daily life: A systematic review and meta-analysis of mnemonic strategy, cognitive training, external memory aid, and combination interventions. *Neuropsychology*, 35(1), 123–140. <https://doi.org/10.1037/neu0000704>
- Kang, E. (2022). Easily accessible but easily forgettable: How ease of access to information online affects cognitive miserliness. *Journal of Experimental*

- Psychology. Applied*. <https://doi.org/10.1037/xap0000412>
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–4311. <https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997>
- Keijzer, F. (2002). Representation in dynamical and embodied cognition. *Cognitive Systems Research*, *3*(3), 275–288. [https://doi.org/10.1016/S1389-0417\(02\)00043-8](https://doi.org/10.1016/S1389-0417(02)00043-8)
- Kelly, A. J., Perdue, B. M., Love, M. W., Parrish, A. E., & Beran, M. J. (2018). An investigation of prospective memory with output monitoring in preschool children. *The American Journal of Psychology*, *131*(2), 201–210. <https://doi.org/10.5406/amerjpsyc.131.2.0201>
- Kirk, P. A., Robinson, O. J., & Gilbert, S. J. (2021). Trait anxiety does not correlate with metacognitive confidence or reminder usage in a delayed intentions task. *Quarterly Journal of Experimental Psychology*, *74*(4), 634–644. <https://doi.org/10.1177/1747021820970156>
- Kizilirmak, J. M., Rösler, F., Bien, S., & Khader, P. H. (2015). Inferior parietal and right frontal contributions to trial-by-trial adaptations of attention to memory. *Brain Research*, *1614*, 14–27. <https://doi.org/10.1016/j.brainres.2015.04.012>
- Kliegel, M., Ballhausen, N., Hering, A., Ihle, A., Schnitzspahn, K. M., & Zuber, S. (2016). Prospective memory in older adults: Where we are now and what is next. *Gerontology*, *62*, 459–466. <https://doi.org/10.1159/000443698>
- Kliegel, M., & Jäger, T. (2007). The effects of age and cue-action reminders on event-based prospective memory performance in preschoolers. *Cognitive Development*, *22*(1), 33–46. <https://doi.org/10.1016/j.cogdev.2006.08.003>
- 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*(1), 203–208. <https://doi.org/10.1037/0882-7974.23.1.203>
- Kliegel, M., Mackinlay, R., & Jäger, T. (2008). Complex prospective memory: Development across the lifespan and the role of task interruption. *Developmental Psychology*, *44*(2), 612–617. <https://doi.org/10.1037/0012-1649.44.2.612>
- Kliegel, M., Mahy, C. E. V., Voigt, B., Henry, J. D., Rendell, P. G., & Aberle, I. (2013). The development of prospective memory in young schoolchildren: The impact of ongoing task absorption, cue salience, and cue centrality. *Journal of Experimental Child Psychology*, *116*(4), 792–810. <https://doi.org/10.1016/j.jecp.2013.07.012>

- Kliegel, M., & Martin, M. (2003). Prospective memory research: Why is it relevant? *International Journal of Psychology, 38*(4), 193–194.
- Kliegel, O., Pastötter, B., & Bäuml, K.-H. T. (2013). List-method directed forgetting can be selective: Evidence from the 3-list and the 2-list tasks. *Memory & Cognition, 41*(3), 452–464. <https://doi.org/10.3758/s13421-012-0264-7>
- Kliegel, O., Pastötter, B., & Bäuml, K.-H. T. (2020). Does amount of pre-cue encoding modulate selective list method directed forgetting? *Frontiers in Psychology, 11*, 1403–1403. <https://doi.org/10.3389/fpsyg.2020.01403>
- Knight, R. G., Harnett, M., & Titov, N. (2005). The effects of traumatic brain injury on the predicted and actual performance of a test of prospective remembering. *Brain Injury, 19*(1), 19–27. <https://doi.org/10.1080/02699050410001720022>
- Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General, 139*(4), 665–682. <https://doi.org/10.1037/a0020198>
- Koriat, A. (2007). *Metacognition and consciousness*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511816789.012>
- Kuhlmann, B. G. (2019). Metacognition of prospective memory. Will I remember to remember? In J. Rummel & M. A. McDaniel (Eds.), *Prospective Memory* (pp. 5–18). Routledge.
- Kurzban, R. (2016). The sense of effort. *Current Opinion in Psychology, 7*, 67–70.
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *The Behavioral and Brain Sciences, 36*(6), 661–679. <https://doi.org/10.1017/S0140525X12003196>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13 SE-Articles), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- 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*(1), 112–132. <https://doi.org/10.1037/0096-3445.136.1.112>
- Lakshmi, A., Wittenbrink, B., Correll, J., & Ma, D. S. (2021). The India face set: International and cultural boundaries impact face impressions and perceptions of category membership. *Frontiers in Psychology, 12*, 627678–627678. <https://doi.org/10.3389/fpsyg.2021.627678>
- Lemaire, P., & Lecacheur, M. (2010). Strategy switch costs in arithmetic problem solving. *Memory & Cognition, 38*(3), 322–332. <https://doi.org/10.3758/MC.38.3.322>
- Levy, B. J., & Anderson, M. C. (2012). Purging of memories from conscious awareness

- tracked in the human brain. *The Journal of Neuroscience*, 32(47), 16785–16794. <https://doi.org/10.1523/JNEUROSCI.2640-12.2012>
- Lewis-Peacock, J. A., & Norman, K. A. (2014). Competition between items in working memory leads to forgetting. *Nature Communications*, 5(1), 5768–5768. <https://doi.org/10.1038/ncomms6768>
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Journal of experimental social psychology detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Experimental Social Psychology*, 4–6.
- Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance*, 20(2), 159–183. [https://doi.org/10.1016/0030-5073\(77\)90001-0](https://doi.org/10.1016/0030-5073(77)90001-0)
- Lindenberger, U., & Mayr, U. (2013). Cognitive aging: Is there a dark side to environmental support? *Trends in Cognitive Sciences*, 18(1), 7–15. <https://doi.org/10.1016/j.tics.2013.10.006>
- Lineweaver, T. T., Horhota, M., Crumley, J., Geanon, C. T., & Juett, J. J. (2018). Age differences in perceptions of memory strategy effectiveness for recent and remote memory. *Aging, Neuropsychology, and Cognition*, 25(2), 146–166. <https://doi.org/10.1080/13825585.2016.1269146>
- 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(3), 441–456. <https://doi.org/10.1080/09541440340000114>
- Lourenço, J. S., & Maylor, E. A. (2015). When distraction holds relevance: A prospective memory benefit for older adults. *International Journal of Environmental Research and Public Health*, 12(6), 6523–6541. <https://doi.org/10.3390/ijerph120606523>
- Lovelace, E. a., & Twohig, P. T. (1990). Healthy older adult's perceptions of their memory functioning and use of mnemonics. *Bulletin of the Psychonomic Society*, 28(2), 115–118. <https://doi.org/10.3758/BF03333979>
- Luis, C. A., Keegan, A. P., & Mullan, M. (2009). Cross validation of the Montreal Cognitive Assessment in community dwelling older adults residing in the Southeastern US. *International Journal of Geriatric Psychiatry*, 24(2), 197–201. <https://doi.org/10.1002/gps.2101>
- Lurie, R., & Westerman, D. L. (2021). Photo-taking impairs memory on perceptual and conceptual memory tests. *Journal of Applied Research in Memory and Cognition*, 10(2), 289–297. <https://doi.org/10.1016/j.jarmac.2020.11.002>

- Lurquin, J. H., & Miyake, A. (2017). Challenges to ego-depletion research go beyond the replication crisis: A need for tackling the conceptual crisis. *Frontiers in Psychology, 8*(APR), 1–5. <https://doi.org/10.3389/fpsyg.2017.00568>
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*(4), 1122–1135. <https://doi.org/10.3758/s13428-014-0532-5>
- Ma, D. S., Kantner, J., & Wittenbrink, B. (2020). Chicago face database: Multiracial expansion. *Behavior Research Methods, 53*(3), 1289–1300. <https://doi.org/10.3758/s13428-020-01482-5>
- Mahy, C. E. V., Mazachowsky, T. R., & Pagobo, J. R. (2018). Do verbal reminders improve preschoolers' prospective memory performance? It depends on age and individual differences. *Cognitive Development, 47*, 158–167. <https://doi.org/10.1016/j.cogdev.2018.06.004>
- Manly, T., Hawkins, K., Evans, J., Woldt, K., & Robertson, I. H. (2002). Rehabilitation of executive function: Facilitation of effective goal management on complex tasks using periodic auditory alerts. *Neuropsychologia, 40*(3), 271–281. [https://doi.org/10.1016/S0028-3932\(01\)00094-X](https://doi.org/10.1016/S0028-3932(01)00094-X)
- Manning, J. R., Hulbert, J. C., Williams, J., Piloto, L., Sahakyan, L., & Norman, K. A. (2016). A neural signature of contextually mediated intentional forgetting. *Psychonomic Bulletin & Review, 23*(5), 1534–1542. <https://doi.org/10.3758/s13423-016-1024-7>
- Mäntylä, T., Missier, F. Del, & Nilsson, L.-G. (2009). Age differences in multiple outcome measures of time-based prospective memory. *Aging, Neuropsychology and Cognition, 16*(6), 708–720.
- Marsh, E. J., & Rajaram, S. (2019). The digital expansion of the mind: implications of internet usage for memory and cognition. *Journal of Applied Research in Memory and Cognition, 8*(1), 1–14. <https://doi.org/10.1016/j.jarmac.2018.11.001>
- Mary, A., Dayan, J., Leone, G., Postel, C., Fraisse, F., Malle, C., Vallée, T., Klein-Peschanski, C., Viader, F., de la Sayette, V., Peschanski, D., Eustache, F., & Gagnepain, P. (2020). Resilience after trauma: The role of memory suppression. *Science (American Association for the Advancement of Science), 367*(6479), 756. <https://doi.org/10.1126/science.aay8477>
- Maylor, E. A. (1990). Age and prospective memory. In *THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY* (Vol. 42, Issue 3).
- Maylor, E. A. (1995). Prospective memory in normal ageing and dementia. *Neurocase, 1*(3), 285–289. <https://doi.org/10.1080/13554799508402372>
- McCurdy, L. Y., Maniscalco, B., Metcalfe, J., Liu, K. Y., de Lange, F. P., & Lau, H.

- (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *Journal of Neuroscience*, *33*(5), 1897–1906. <https://doi.org/10.1523/JNEUROSCI.1890-12.2013>
- McDaniel, M. A., Bugg, J. M., Ramuschkat, G. M., Kliegel, M., & Einstein, G. O. (2009). Repetition errors in habitual prospective memory: Elimination of age differences via complex actions or appropriate resource allocation. *Aging, Neuropsychology, and Cognition*, *16*(5), 563–588. <https://doi.org/10.1080/13825580902866646>
- McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology*, *14*(7), S127–S144. <https://doi.org/10.1002/acp.775>
- McDaniel, M. A., & Einstein, G. O. (2011). The neuropsychology of prospective memory in normal aging: A componential approach. *Neuropsychologia*, *49*(8), 2147–2155. <https://doi.org/10.1016/j.neuropsychologia.2010.12.029>
- McDaniel, M. A., Guynn, M. J., Einstein, G. O., & Breneiser, J. (2004). Cue-focused and reflexive-associative processes in prospective memory retrieval. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(3), 605–614. <https://doi.org/10.1037/0278-7393.30.3.605>
- McDaniel, M. A., Robinson-Riegler, B., & Einstein, G. O. (1998). Prospective remembering: Perceptually driven or conceptually driven processes? *Memory & Cognition*, *26*(1), 121–134. <https://doi.org/10.3758/BF03211375>
- McDaniel, M. A., Umanath, S., Einstein, G. O., & Waldum, E. R. (2015). Dual pathways to prospective remembering. *Frontiers in Human Neuroscience*, *9*, 392–392. <https://doi.org/10.3389/fnhum.2015.00392>
- McGillivray, S., & Castel, A. D. (2017). OLDER AND YOUNGER ADULTS' STRATEGIC CONTROL OF METACOGNITIVE MONITORING: THE ROLE OF CONSEQUENCES, TASK EXPERIENCE, AND PRIOR KNOWLEDGE. *Experimental Aging Research*, *43*(3), 233–256. <https://doi.org/10.1080/0361073X.2017.1298956>
- Meeks, J. T, Hicks, J. L., & Marsh, R. L. (2007). Metacognitive awareness of event-based prospective memory. *Consciousness and Cognition*, *16*(4), 997–1004. <https://doi.org/10.1016/j.concog.2006.09.005>
- Mioni, G., Grondin, S., McLennan, S. N., & Stablum, F. (2020). The role of time-monitoring behaviour in time-based prospective memory performance in younger and older adults. *Memory*, *28*(1), 34–48. <https://doi.org/10.1080/09658211.2019.1675711>
- Mioni, G., & Stablum, F. (2014). Monitoring behaviour in a time-based prospective memory task: The involvement of executive functions and time perception. *Memory*, *22*(5), 536–552. <https://doi.org/10.1080/09658211.2013.801987>

- Morrison, A. B., & Richmond, L. L. (2020). Offloading items from memory: Individual differences in cognitive offloading in a short-term memory task. *Cognitive Research: Principles and Implications*, 5(1), 1–13. <https://doi.org/10.1186/s41235-019-0201-4>
- Moscovitch, M. (1982). A neuropsychological approach to perception and memory in normal and pathological aging. In *Aging and cognitive processes* (pp. 55–78). Springer US. https://doi.org/10.1007/978-1-4684-4178-9_4
- Mruczek, R. E. B., Killebrew, K. W., & Berryhill, M. E. (2019). Individual differences reveal limited mixed-category effects during a visual working memory task. *Neuropsychologia*, 122, 1–10. <https://doi.org/10.1016/j.neuropsychologia.2018.12.005>
- Murphy, D. H., & Castel, A. D. (2020). Responsible remembering: How metacognition impacts adaptive selective memory. *Zeitschrift Für Psychologie*, 228(4), 301–303. <https://doi.org/10.1027/2151-2604/a000428>
- Murphy, D. H., & Castel, A. D. (2021). Responsible remembering and forgetting as contributors to memory for important information. *Memory & Cognition*, 49(5), 895–911. <https://doi.org/10.3758/s13421-021-01139-4>
- Nee, D. E., & Jonides, J. (2009). Common and distinct neural correlates of perceptual and memorial selection. *NeuroImage (Orlando, Fla.)*, 45(3), 963–975. <https://doi.org/10.1016/j.neuroimage.2009.01.005>
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation*, 26, 125–173. [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
- Nettelbeck, T., & Rabbitt, P. M. A. (1992). Aging, cognitive performance, and mental speed. *Intelligence*, 16(2), 189–205. [https://doi.org/10.1016/0160-2896\(92\)90004-B](https://doi.org/10.1016/0160-2896(92)90004-B)
- Newman, E. L., & Norman, K. A. (2010). Moderate excitation leads to weakening of perceptual representations. *Cerebral Cortex (New York, N.Y. 1991)*, 20(11), 2760–2770. <https://doi.org/10.1093/cercor/bhq021>
- Norman, K. A., L. Newman, E., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review*, 114(4), 887–953. <https://doi.org/10.1037/0033-295X.114.4.887>
- Norton, M. I., Mochon, D., & Ariely, D. (2012). The IKEA effect: When labor leads to love. *Journal of Consumer Psychology*, 22(3), 453–460. <https://doi.org/10.1016/j.jcps.2011.08.002>
- Oehr, C. R., Fell, J., Baumann, C., Rosburg, T., Ludowig, E., Kessler, H., Hanslmayr, S., & Axmacher, N. (2018). Direct electrophysiological evidence for prefrontal control of hippocampal processing during voluntary forgetting. *Current*

- Biology*, 28(18), 3016–3022.e4. <https://doi.org/10.1016/j.cub.2018.07.042>
- Olivola, C. Y. (2018). The interpersonal sunk-cost effect. *Psychological Science*, 29(7), 1072–1083. <https://doi.org/10.1177/0956797617752641>
- Oostrom, J. K., van der Linden, D., Born, M. Ph., & van der Molen, H. T. (2013). New technology in personnel selection: How recruiter characteristics affect the adoption of new selection technology. *Computers in Human Behavior*, 29(6), 2404–2415. <https://doi.org/10.1016/j.chb.2013.05.025>
- Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of Cognitive Neuroscience*, 23(11), 3419–3432. https://doi.org/10.1162/jocn_a_00011
- Palejwala, A. H., Dadario, N. B., Young, I. M., O'Connor, K., Briggs, R. G., Conner, A. K., O'Donoghue, D. L., & Sughrue, M. E. (2021). Anatomy and white matter connections of the lingual gyrus and cuneus. *World Neurosurgery*, 151, e426–e437. <https://doi.org/10.1016/j.wneu.2021.04.050>
- Park, D. C., Hertzog, C., Kidder, D. P., Morrell, R. W., & Mayhorn, C. B. (1997). Effect of age on event-based and time-based prospective memory. *Psychology and Aging*, 12(2), 314–327. <https://doi.org/10.1037//0882-7974.12.2.314>
- Pastötter, B., Kliegl, O., & Bäuml, K.-H. T. (2012). List-method directed forgetting: The forget cue improves both encoding and retrieval of postcue information. *Memory & Cognition*, 40(6), 861–873. <https://doi.org/10.3758/s13421-012-0206-4>
- Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58(3), 193–198. <https://doi.org/10.1037/h0049234>
- 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.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives*. (pp. 161–186). Erlbaum.
- Prospective memory: Theory and applications. (1996). In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications*. Lawrence Erlbaum Associates Publishers.
- R Core, T. (2021). *R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing.*
- Racsmány, M., & Conway, M. A. (2006). Episodic inhibition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 32(1), 44–57. <https://doi.org/10.1037/0278-7393.32.1.44>
- Redshaw, J., Vandersee, J., Bulley, A., & Gilbert, S. J. (2018). Development of children's use of external reminders for hard-to-remember intentions. *Child*

- Development*, 00(0), 1–10. <https://doi.org/10.1111/cdev.13040>
- Rendell, P. G., & Craik, F. I. M. (2000). Virtual week and actual week: Age-related differences in prospective memory. *Applied Cognitive Psychology*, 14(7), S43–S62. <https://doi.org/10.1002/acp.770>
- 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(4), P256–P269. <https://doi.org/10.1093/geronb/54B.4.P256>
- Risko, E. F., & Dunn, T. L. (2015). Storing information in-the-world: Metacognition and cognitive offloading in a short-term memory task. *Consciousness and Cognition*, 36, 61–74. <https://doi.org/10.1016/j.concog.2015.05.014>
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. <https://doi.org/10.1016/j.tics.2016.07.002>
- 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(3), 595–605. <https://doi.org/10.1037/a0019771>
- Rummel, J., Danner, D., & Kuhlmann, B. G. (2019). The short version of the Metacognitive Prospective Memory Inventory (MPMI-s): Factor structure, reliability, validity, and reference data. *Measurement Instruments for the Social Sciences*, 1(1), 1–8. <https://doi.org/10.1186/s42409-019-0008-6>
- Rummel, J., & Meiser, T. (2013). The role of metacognition in prospective memory: Anticipated task demands influence attention allocation strategies. *Consciousness and Cognition*, 22(3), 931–943. <https://doi.org/10.1016/j.concog.2013.06.006>
- Runge, Y., Frings, C., & Tempel, T. (2019). Saving-enhanced performance: Saving items after study boosts performance in subsequent cognitively demanding tasks. *Memory (Hove)*, 27(10), 1462–1467. <https://doi.org/10.1080/09658211.2019.1654520>
- Runge, Y., Frings, C., & Tempel, T. (2020). Specifying the mechanisms behind benefits of saving-enhanced memory. *Psychological Research*, 85(4), 1633–1644. <https://doi.org/10.1007/s00426-020-01341-0>
- Runge, Y., Frings, C., Tempel, T., & Pastötter, B. (2021). Electrophysiological correlates of saving-enhanced memory: Exploring similarities to list-method directed forgetting. *The European Journal of Neuroscience*, 54(6), 6060–6074. <https://doi.org/10.1111/ejn.15368>
- Ryan, A. D., & Campbell, K. L. (2021). The ironic effect of older adults’ increased task motivation: Implications for neurocognitive aging. *Psychonomic Bulletin &*

- Review. <https://doi.org/10.3758/s13423-021-01963-4>
- Ryder, N., Kvavilashvili, L., & Ford, R. (2022). Effects of incidental reminders on prospective memory in children. *Developmental Psychology, 58*(5), 890–901. <https://doi.org/10.1037/dev0001035>
- Sachdeva, C., & Gilbert, S. J. (2020). Excessive use of reminders: Metacognition and effort-minimisation in cognitive offloading. *Consciousness and Cognition, 85*(September), 1–14. <https://doi.org/10.1016/j.concog.2020.103024>
- Sahakyan, L., & Delaney, P. F. (2005). Directed forgetting in incidental learning and recognition testing: Support for a two-factor account. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*(4), 789–801. <https://doi.org/10.1037/0278-7393.31.4.789>
- Sahakyan, L., Delaney, P. F., Foster, N. L., & Abushanab, B. (2013). Chapter four—List-method directed forgetting in cognitive and clinical research: A theoretical and methodological review. *The Psychology of Learning and Motivation, 59*, 131–189.
- Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 28*(6), 1064–1072.
- Scarampi, C., & Gilbert, S. J. (2020). The effect of recent reminder setting on subsequent strategy and performance in a prospective memory task. *Memory, 28*(5), 677–691. <https://doi.org/10.1080/09658211.2020.1764974>
- Scarampi, C., & Gilbert, S. J. (2021). Age differences in strategic reminder setting and the compensatory role of metacognition. *Psychology and Aging, 36*(2), 172–185. <https://doi.org/10.1037/pag0000590>
- Scarampi, C., & Kliegel, M. (2021). *Metamemory for prospective memory performance in younger and older adults: Does the reference point affect our judgments?* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/56kdv>
- Schmitz, T. W., Correia, M. M., Ferreira, C. S., Prescott, A. P., & Anderson, M. C. (2017). Hippocampal GABA enables inhibitory control over unwanted thoughts. *Nature Communications, 8*(1), 1311–1312. <https://doi.org/10.1038/s41467-017-00956-z>
- 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*(4), 583–592. <https://doi.org/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*(8), 1544–1553.

- <https://doi.org/10.1037/a0030579>
- Schnitzspahn, K. M., Zeintl, M., Jäger, T., & Kliegel, M. (2011). Metacognition in prospective memory: Are performance predictions accurate? *Canadian Journal of Experimental Psychology, 65*(1), 19–26.
<https://doi.org/10.1037/a0022842>
- Scullin, M. K., McDaniel, M. A., Dasse, M. N., Lee, J. H., Kurinec, C. A., Tami, C., & Krueger, M. L. (2018). Thought probes during prospective memory encoding: Evidence for perfunctory processes. *PLoS One, 13*(6), e0198646–e0198646.
<https://doi.org/10.1371/journal.pone.0198646>
- Scullin, M. K., McDaniel, M. A., & Shelton, J. T. (2013). The dynamic multiprocess framework: Evidence from prospective memory with contextual variability. *Cognitive Psychology, 67*(1–2), 55–71.
<https://doi.org/10.1016/j.cogpsych.2013.07.001>
- Shah, J., Higgins, T., & Friedman, R. S. (1998). Performance incentives and means: How regulatory focus influences goal attainment. *Journal of Personality and Social Psychology, 74*(2), 285–293. <https://doi.org/10.1037/0022-3514.74.2.285>
- 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 : A Journal of the American Psychological Society, 26*(4), 352–358. <https://doi.org/10.1177/0963721417700504>
- Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience, 19*(1), 83–90.
<https://doi.org/10.1146/annurev-neuro-072116-031526>
- Sheppard, D. P., Matchanova, A., Sullivan, K. L., Kazimi, S. I., & Woods, S. P. (2020). Prospective memory partially mediates the association between aging and everyday functioning. *The Clinical Neuropsychologist, 34*(4), 755–774.
<https://doi.org/10.1080/13854046.2019.1637461>
- Shum, D. (2004). Effects of incentive and preparation time on performance and classification accuracy of standard and malingering-specific memory tests. *Archives of Clinical Neuropsychology, 19*(6), 817–823.
<https://doi.org/10.1016/j.acn.2003.10.002>
- Siegel, A. L. M., & Castel, A. D. (2019). Age-related differences in metacognition for memory capacity and selectivity. *Memory, 27*(9), 1236–1249.
<https://doi.org/10.1080/09658211.2019.1645859>
- Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *The Behavioral and Brain*

- Sciences*, 26(3), 317–339. <https://doi.org/10.1017/S0140525X03000086>
- 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(3), 347–361. <https://doi.org/10.1037/0278-7393.29.3.347>
- Smith, R. E., Bayen, U. J., & Martin, C. (2010). The cognitive processes underlying event-based prospective memory in school-age children and young adults: A formal model-based study. *Developmental Psychology*, 46(1), 230–244. <https://doi.org/10.1037/a0017100>
- Smith, R. E., Horn, S. S., & Bayen, U. J. (2012). Prospective memory in young and older adults: The effects of ongoing-task load. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 19(4), 495–514. <https://doi.org/10.1080/13825585.2011.633161> [doi]
- Smith, R. E., Hunt, R. R., & Murray, A. E. (2017). Prospective memory in context: Moving through a familiar space. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 43(2), 189–204. <https://doi.org/10.1037/xlm0000303>
- Smith, R. E., & Loft, S. (2014). Investigating the cost to ongoing tasks not associated with prospective memory task requirements. *Consciousness and Cognition*, 27, 1–13. <https://doi.org/10.1016/j.concog.2014.04.002>
- Smith, R. E., & Skinner, D. M. (2019). Prospective memory in context: Methods, findings, and future directions. In J. Rummel & M. A. McDaniel (Eds.), *Prospective Memory*. Routledge.
- Soderstrom, N. C., McCabe, D. P., & Rhodes, M. G. (2012). Older adults predict more recollective experiences than younger adults. *Psychology and Aging*, 27(4), 1082–1088. <https://doi.org/10.1037/a0029048>
- Souchay, C. (2007). Metamemory in Alzheimer’s disease. *Cortex*, 43(7), 987–1003. [https://doi.org/10.1016/S0010-9452\(08\)70696-8](https://doi.org/10.1016/S0010-9452(08)70696-8)
- Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science (American Association for the Advancement of Science)*, 333(6043), 776–778. <https://doi.org/10.1126/science.1207745>
- Spitzer, B. (2014). Finding retrieval-induced forgetting in recognition tests: A case for baseline memory strength. *Frontiers in Psychology*, 5, 1102–1102. <https://doi.org/10.3389/fpsyg.2014.01102>
- Stevens, M. C., Calhoun, V. D., & Kiehl, K. A. (2005). Hemispheric differences in

- hemodynamics elicited by auditory oddball stimuli. *NeuroImage (Orlando, Fla.)*, 26(3), 782–792. <https://doi.org/10.1016/j.neuroimage.2005.02.044>
- Storm, B. C., & Stone, S. M. (2015). Saving-enhanced memory: The benefits of saving on the learning and remembering of new information. *Psychological Science*, 26(2), 182–188. <https://doi.org/10.1177/0956797614559285>
- Strayer, D. L., & Kramer, A. F. (1994). Strategies and automaticity: I. Basic findings and conceptual framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(2), 318–341. <https://doi.org/10.1037/0278-7393.20.2.318>
- Suija, K., Rajala, U., Jokelainen, J., Liukkonen, T., Härkönen, P., Keinänen-Kiukaanniemi, S., & Timonen, M. (2012). Validation of the Whooley questions and the Beck Depression Inventory in older adults. *Scandinavian Journal of Primary Health Care*, 30(4), 259–264. <https://doi.org/10.3109/02813432.2012.732473>
- Taillan, J., Ardiale, E., & Lemaire, P. (2015). Relationships between strategy switching and strategy switch costs in young and older adults: A study in arithmetic problem solving. *Experimental Aging Research*, 41(2), 136–156. <https://doi.org/10.1080/0361073X.2015.1001651>
- Tan, P., Ensor, T. M., Hockley, W. E., Harrison, G. W., & Wilson, D. E. (2020). In support of selective rehearsal: Double-item presentation in item-method directed forgetting. *Psychonomic Bulletin & Review*, 27(3), 529–535. <https://doi.org/10.3758/s13423-020-01723-w>
- Tisserand, D. J., van Boxtel, M. P. J., Pruessner, J. C., Hofman, P., Evans, A. C., & Jolles, J. (2004). A voxel-based morphometric study to determine individual differences in gray matter density associated with age and cognitive change over time. *Cerebral Cortex (New York, N.Y. 1991)*, 14(9), 966–973. <https://doi.org/10.1093/cercor/bhh057>
- Touron, D. R. (2015). Memory avoidance by older adults: When ‘old dogs’ won’t perform their ‘new tricks’. *Current Directions in Psychological Science*, 24(3), 170–176. <https://doi.org/10.1177/0963721414563730>
- Touron, D. R., & Hertzog, C. (2014). Accuracy and speed feedback: Global and local effects on strategy use. *Experimental Aging Research*, 40(3), 332–356. <https://doi.org/10.1080/0361073X.2014.897150>
- Touron, D. R., Swaim, E. T., & Hertzog, C. (2007). Moderation of older adults’ retrieval reluctance through task instructions and monetary incentives. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 62(3), P149–P155. <https://doi.org/10.1093/geronb/62.3.P149>
- Uretzky, S., & Gilboa, A. (2010). Knowing your lines but missing your cue: rostral

- prefrontal lesions impair prospective memory cue detection, but not action-intention superiority. *Journal of Cognitive Neuroscience*, 22(12), 2745–2757. <https://doi.org/10.1162/jocn.2010.21419>
- Uttl, B. (2008). Transparent meta-analysis of prospective memory and aging. *PLoS ONE*, 3(2). <https://doi.org/10.1371/journal.pone.0001568>
- Vossel, S., Thiel, C. M., & Fink, G. R. (2006). Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. *NeuroImage (Orlando, Fla.)*, 32(3), 1257–1264. <https://doi.org/10.1016/j.neuroimage.2006.05.019>
- Waldum, E. R., & McDaniel, M. A. (2016). Why are you late? Investigating the role of time management in time-based prospective memory. *Journal of Experimental Psychology. General*, 145(8), 1049–1061. <https://doi.org/10.1037/xge0000183>
- Wang, L., Baumgartner, F., Kaule, F. R., Hanke, M., & Pollmann, S. (2019). Individual face- and house-related eye movement patterns distinctively activate FFA and PPA. *Nature Communications*, 10(1), 5532–16. <https://doi.org/10.1038/s41467-019-13541-3>
- Ward, A. F. (2013). Supernormal: How the internet is changing our memories and our minds. *Psychological Inquiry*, 24(4), 341–348. <https://doi.org/10.1080/1047840X.2013.850148>
- Weis, P. P., & Wiese, E. (2019). Using tools to help us think: Actual but also believed reliability modulates cognitive offloading. *Human Factors*, 61(2), 243–254. <https://doi.org/10.1177/0018720818797553>
- Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PLoS ONE*, 8(7), 1–8. <https://doi.org/10.1371/journal.pone.0068210>
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9(4), 625–636. <https://doi.org/10.3758/BF03196322>
- Zacks, R. T., Radvansky, G., & Hasher, L. (1996). Studies of directed forgetting in older adults. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 22(1), 143–156. <https://doi.org/10.1037/0278-7393.22.1.143>
- Zahodne, L. B., & Reuter-Lorenz, P. A. (2019). COMPENSATION AND BRAIN AGING: A REVIEW AND ANALYSIS OF EVIDENCE. In G. R. Samanez-Larkin (Ed.), *The Aging Brain: Functional Adaptation across Adulthood* (pp. 185–216). American Psychological Association. <https://doi.org/10.1037/0000143-008>
- Zuber, S., & Kliegel, M. (2020). Prospective memory development across the lifespan: An integrative framework. *European Psychologist*, 25(3), 162–173.

<https://doi.org/10.1027/1016-9040/a000380>