CAUSES AND CONSEQUENCES OF COGNITIVE OFFLOADING

Chhavi Sachdeva

Student Number: 14067902

University College London

Institute of Cognitive Neuroscience

Faculty of Brain Sciences

Dissertation submitted for the degree of

Doctor of Philosophy in Cognitive Neuroscience

November 2022

I, *Chhavi Sachdeva*, 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:* 11th November, 2022

ABSTRACT

The current thesis focuses on cognitive offloading. The first three chapters explore factors influencing cognitive offloading, namely metacognition and effort-minimisation while the last chapter focuses on the consequences of cognitive offloading on subsequently remembered information. The first chapter investigated whether metacognitive interventions designed to shift confidence also influence offloading behaviour. It was found that interventions designed to shift confidence also shifted participants' offloading behaviour. It was also found that confidence cannot fully explain offloading behaviour. The second chapter explored whether other factors such as preference to avoid cognitive effort contribute to offloading behaviour. It was found that this factor influenced offloading such that the bias towards offloading was reduced (but not eliminated) in the group that received performance-based rewards, hypothesised to reduce effort-avoidance. The third chapter sought to examine whether offloading behaviour was also related to confidence in a task from an unrelated domain (in this case a pair of perceptual tasks). This chapter found that perceptual confidence was related with propensity to offload but not preference to offload, relative to the optimal strategy. The final chapter focused on the consequences of offloading where in the first experiment it was found that saving a list of words not only improved memory of that list but also improved memory for subsequently encoded information. However, this was dependent on the order in which the two lists were tested. The second experiment found that participants had a preference towards list-saving in a manner that matched the optimal strategy demonstrated by the first experiment. Collectively, the findings of this thesis will help our understanding of cognitive offloading so that we can guide individuals towards more effective offloading strategies to supplement memory. (277 words).

IMPACT STATEMENT

With the rapid technological advancement of today’s world, cognitive offloading has become an increasingly ubiquitous phenomenon. Despite its pervasiveness, research in cognitive offloading is still in its infancy. The current thesis aimed to explore the phenomenon of cognitive offloading, particularly the reasons why individuals might choose to offload and the consequences of doing so. This thesis makes three core contributions within academia.

First, this thesis emphasises to the reader that cognitive offloading is influenced by various factors. Furthermore, cognitive offloading is multifaceted where what might influence one domain of cognitive offloading might not influence another. Therefore, future research must consider disentangling the processes involved in cognitive offloading in different domains. This in turn, would aid the examination of conditions under which individuals might choose to offload.

Second, this thesis has shown that interventions designed to shift confidence can also influence individuals’ preference towards using reminders. Therefore, it becomes necessary to study metacognitive interventions that influence offloading behaviour as training individuals’ confidence could guide their use of reminders. Relatedly, I also show that domain-general confidence signals play a role in influencing offloading behaviour where propensity to offload is associated with confidence in an unrelated domain. Future research should also aim to elucidate the domain-general versus task-specific signals between the different offloading domains as this will aid in the development of real-world metacognitive interventions to improve individuals’ cognitive offloading strategies.

Third, this thesis addresses the consequences of cognitive offloading. Given the increasing prevalence of cognitive offloading in today’s world, scepticism towards relying on technology has become widepsread. Instead of arguing that supplementing our cognition with external tools is detrimental to our internal cognitive processes, the current thesis argues that perhaps supplementing our cognition is, in some cases, *beneficial* to our internal cognitive processes. This is argued by showing that saving material frees cognitive resources such that these can then be reallocated towards other tasks where when given the opportunity to save a list of words, recall performance not only increases for that list of words but also for subsequently presented material. Future research should, therefore, investigate cognitive offloading as a tool to improve our daily organisation as opposed to something that might be detrimental to our cognition.

Turning our attention to the real world, cognitive offloading could be an effective tool for supplementing everyday cognition. This is especially true when it comes to older adults and individuals with brain injury. Fulfilling delayed intentions is important to living an independent life. Therefore, developing tools to effectively guide individuals to use external tools will improve independence and ability to complete everyday tasks. The current thesis encourages the use of intervention training to optimise individuals' use of external tools to guide behaviour and raises the importance of future research in this area to fully understand the reasons why an individual might offload and the consequences of doing so. (473 words).

ACKNOWLEDGEMENTS

First, I would like to thank my supervisor, Prof. Sam Gilbert for being an excellent supervisor and giving me the support and opportunity to pursue my research question. I am extremely grateful to not only have had him as an excellent supervisor but also an incredible teacher whom I have learned much from. I would also like to thank my fellow lab mates, Dr. Annika Boldt, Dr. Chiara Scarampi and Pei-Chun Tsai for providing me with a lot of support and wisdom during my PhD years. Above all, I am grateful for their friendship, and I consider myself incredibly lucky to have crossed paths with them.

I would also like to thank Tris for all those hours of patient listening (and counselling) while indulging on Biang Biang noodles and to Kerstin and Marina who have been a huge part of my PhD journey. I would also like to thank Alex who has not only been a huge source of moral support throughout my PhD, but has also been someone I could lean on. Our countless nights sourcing the best Chinese and Asian food reminiscent to what we get in our hometown, Hong Kong, has been a huge source of comfort. On a personal note, a huge shout out to my parents without whom this thesis would not exist. Thank you for your unwavering support, your belief in me and for your hours of patient listening on the phone while we try to wrangle with time zones. I am incredibly grateful for everything you have done for me and to have parents such as you – I dedicate this work to you both.

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

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]: ☒ *I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.*		
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. Chhavi Sachdeva was involved with the conceptualisation and data analysis of experiment 3.

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:	Chhavi Sachdeva	**Date:**	04/11/2022
Supervisor/ Senior Author (where appropriate)**:**	Sam Gilbert	**Date:**	08/11/2022

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

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)	Consciousness and Cognition
g) Who published the work? (e.g. Elsevier/Oxford University Press):	Elsevier
h) When was the work published?	2020
i) Was the work subject to academic peer review?	Yes
j) 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]: ☒ *I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.*	

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):	Please select.	**If yes, which server?** Click or tap here to enter text.
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.	
g) List the manuscript's authors in the intended authorship order:	Click or tap here to enter text.	
h) Stage of publication	Please select.	

8. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4**):**			
Equal contribution to all aspects.			
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:	Chhavi Sachdeva	**Date:**	04/11/2022
Supervisor/ Senior Author (where appropriate)**:**	Sam Gilbert	**Date:**	08/11/2022

Table of Contents

CHAPTER 1. INTRODUCTION

Outline

Imagine a scenario where your doctor prescribes you with some medication that needs to be taken at regular intervals throughout the day. You could choose to remember this by maintaining the intention internally. Or instead, you could create an external reminder by placing the medication next to your bed, setting an alarm on your phone, or asking “Siri” or “Alexa” to remind you. The latter are all examples of cognitive offloading, i.e., the use of physical action to reduce the cognitive demands of a task (Risko & Gilbert, 2016). Due to the rapid development of technology, the use of cognitive offloading to supplement memory has become increasingly ubiquitous. Therefore, it is important to understand when individuals might choose to rely on these external resources and what the consequences are of doing so. This is so that individuals can be guided towards more effective use of external tools to facilitate remembering.

This thesis will be divided into four empirical chapters where the first three empirical chapters will focus on factors influencing cognitive offloading and the last empirical chapter will examine the consequences of cognitive offloading.

In the introduction of this thesis, I will first focus on a broad overview of cognitive offloading before discussing the ways in which individuals offload cognition into the environment. I will then consider early paradigms investigating cognitive offloading before detailing more recent paradigms examining intention offloading which is more specific to this thesis. I will then give an overview of the role of various predictors (more specifically, metacognition, individual differences, and strategy preservation) in cognitive offloading before detailing the consequences of offloading for subsequently encoded intentions. Lastly, I will present a brief overview of the empirical chapters and the research questions they aim to answer.

Cognitive offloading

Offloading our cognition by manipulating our bodies and objects in the external environment is a process that we engage in every day. For example, we physically tilt our heads to perceive rotated images (e.g., Dunn & Risko, 2016; Risko et al., 2014), we use a GPS device for navigation (e.g., Brügger et al., 2019), we use computers (e.g., Storm & Stone, 2015) or notes (e.g., Holbrook & Dismukes, 2009; Kelly & Risko, 2019; Marsh et al., 1998) to remember delayed intentions, we use photographs to aid memory (e.g., Henkel, 2014; Soares & Storm, 2018), or we use the internet to look for information rather than use our own internal memory (e.g., Marsh & Rajaram, 2019; Sparrow et al., 2011). This complex relationship between our cognition and the environment means that we often think using our bodies and the physical world around us.

Broadly, cognitive offloading can be categorized into actions that offload cognition onto the body and actions that offload cognition into the world (Risko & Gilbert, 2016). When we offload cognition onto our bodies, we physically use our bodies to reduce cognitive demand. Examples of this include tilting your head to perceive rotated images (e.g., Dunn & Risko, 2016; Risko et al., 2014) and using your fingers to simplify arithmetic tasks (e.g., Goldin-Meadow et al., 2001). Offloading cognition into the world involves using the external environment to support cognition (Risko & Gilbert, 2016). For instance, placing post-it notes to support memory or co-operating with others to reduce cognitive interference in task (Tufft & Richardson, 2020). This thesis will mostly focus on instances where we offload cognition into the world.

Practically, cognitive offloading is an effective strategy when it comes to supplementing our memory. We know that human memory is fallible (Cowan, 2010; Gilchrist et al., 2008; Miller, 1956) and what can be encoded or retrieved at any given time is limited (Schacter, 2001). So taking advantage of external resources to enhance cognition can radically expand our cognitive abilities (see Clark & Chalmers, 1998; B. Tversky, 2011) by overcoming these capacity limitations and minimizing cognitive effort. Indeed, cognitive offloading has been shown to improve performance in various domains such as memory for delayed intentions (e.g., Gilbert, 2015a, 2015b; Gilbert et al., 2020), working memory (e.g., Ballard et al., 1992, 1995), perception (e.g., Dunn & Risko, 2016; Risko et al., 2014), math (e.g., Goldin-Meadow et al., 2001), and learning (e.g., Hu et al., 2019).

The notion of cognitive offloading is not new and has been referred to by other terms in cognitive research. For example, Kirsh and Maglio (1994) used the term *epistemic actions* to describe physical, external actions that are performed by an individual to change their own internal computational state, thereby making cognitive computations easier, faster, and more reliable. Similarly, Scaife and Rogers (1996) used the term *computational offloading* to describe the use of graphical representations (such as diagrams, animations and multimedia) in understanding graphical technology. However, until recently, the phenomenon of cognitive offloading was rarely the topic of experimental research. Interest in this area of research has increased because offloading offers a deeper understanding in the distributed nature of human cognition and its consequences in our everyday lives (Risko & Gilbert, 2016). Cognitive offloading also provides an avenue to study how human cognition is supplemented by its environment in terms of the extended mind, embodiment and distributed cognition (Clark, 2010; Clark & Chalmers, 1998; Hutchins, 1995; Norman, 1993; M. Wilson, 2002).

Given this distributed nature of human cognition and the pervasiveness of cognitive offloading in everyday life, it becomes important to understand the mechanisms of cognitive offloading to fully understand the supplementation of human cognition onto external resources.

Offloading onto others

For much of recorded history, humans have offloaded memory tasks (Nestojko et al., 2013). In fact, such offloading is why we have records of history. The ancient Greeks developed and perfected mnemonic systems that helped individuals encode, store and retrieve large amounts of information (Yates, 1966). In ancient Rome, slaves were trained to remember information about social and legal issues where their job was to help their masters when they needed information during speeches or debates (Popkin & Ng, 2022). This kind of offloading is an example of *socially distributed cognition* (Hutchins, 1995).

As in the examples outlined above, socially distributed cognition is the offloading of knowledge or information on to others. In this type of transactive memory system, knowledge is distributed across two or more individuals such that the system as a whole knows more than any one individual (Wegner, 1995). Tufft and Richardson (2020) extended this notion of a transactive memory process to one where socially distributed cognition describes any shared task-based situation where individuals are able to leverage others social agents to facilitate their own cognitive performance in a task. They propose that in appropriate social contexts this might trigger offloading behaviours which might include freeing up cognitive resources through sharing cognitive demands with other agents or increasing the efficiency of ongoing cognitive processes through socially led modulation of cognitive interference (Tufft & Richardson, 2020).

Studies utilising interference paradigms have found that working with a partner reduces interference in these tasks (e.g., see Heed et al., 2010 using the Stroop task; Sharma et al., 2010 using a variation of the Simon task; Tufft & Richardson, 2020 using a Picture Word Interference task). Termed *social offloading* (see Tufft & Richardson, 2020), the view here is that offloading onto another person reduces internal cognitive conflict (i.e., interference) by decreasing the influence of distracting information, thus improving performance. This notion extends on transactive memory in that dividing encoding responsibilities between individuals not only increases the number of items to be remembered by the group (rather than by any one person) but also allows each individual within this system to improve individual performance when working together compared to when working alone. Social offloading, therefore, describes any shared task situation in which an individual is able to leverage other individuals to facilitate their own cognitive performance. An underlying mechanism of this, might include freeing up cognitive resources through sharing task demands with other individuals. Consistent with this notion, Tufft and Richardson (2020) found that in a Picture Word Interference paradigm, task performance improved and cognitive interference decreased when participants believed they were working together with another individual.

In addition to socially distributing our cognition, we are also able alter the physical environment to expand our abilities (Sparrow & Chatman, 2013). This is particularly true when it comes to externalizing our memory processes.

Offloading intentions onto the environment

A different type of offloading is that of offloading cognition onto the environment. Clark and Chalmers (1998) posited that cognitive operations should be understood as hybrid processes which take place both within the human brain and beyond it. This view was extended by Menary (2010) who implied that research should consider the dynamic interplay between internal and external memory storage instead of studying internal cognitive processes in isolation.

An example of where this is applicable is research prospective memory (PM). PM is of fundamental importance in the development and maintenance of an independent and autonomous life (Cockburn & Smith, 1988) and refers to the ability to form intentions and act on them at an appropriate time (Brandimonte et al., 2014; Einstein & McDaniel, 1990). The majority of self-reported everyday memory failures comprise of PM failures (Kliegel & Martin, 2003). The reason why remembering delayed intentions is particularly difficult is because we are not overtly prompted to retrieve the intention and act at the appropriate time, instead we must “remember to remember” (Ellis, 1996). Therefore, individuals often supplement their memory for delayed intentions by setting external reminders.

Consistent with the extended mind hypothesis (Clark & Chalmers, 1998), individuals interact with their environment in a way that our intentions are stored in a system that extends beyond our brains in the physical environment. For example, instead of relying on internal memory processes to remember these intentions, we also rely on external devices such as diaries, notes, strategically placed objects and increasingly, digital devices such as smartphones or virtual assistant technologies (e.g., Siri and Alexa). In this way, rather than relying on our internal cognitive processes, we reorganize our surroundings to create perceptual triggers to fulfil delayed intentions (Kirsh, 1996). Using external cues to supplement our memory for future intentions is a form of cognitive offloading known as intention offloading.

Early research looking at how individuals can successfully fulfil delayed intentions, has suggested that individuals can enhance their cognition by using external resources.

Mäntylä (1996) suggested that planning activities might influence PM by automatically increasing the activation of the representation of delayed intentions. This in turn could benefit PM by increasing the number of retrieval routes of the intention. Direct evidence for this notion comes from one of the very first studies that incorporated reminders into a PM task. This study was conducted by Meacham and Leiman (1982) who found that providing participants with a coloured tag for their keychains increased remembering for the intention of returning postcards to the experimenter. Diary studies have also found that supplementing memory for delayed intentions using external props can greatly improve one's PM performance in everyday situations through explicit planning (Holbrook & Dismukes, 2009; Marsh et al., 1998). More recently, research has also found that setting reminders reduces cognitive demand and improves memory performance in cognitive tasks (e.g., Gilbert et al., 2020; see Guo et al., 2021 for findings in a time-based PM task; Hu et al., 2019; Risko & Dunn, 2015).

However, there are studies that have proposed that using reminders might not always be beneficial. For example, in a simulated air traffic control task, Vortac et al. (1995) found that continuously presenting participants with the content of intended actions (i.e., name of planes and their destinations) during the retention interval did not improve PM, suggesting that reminders do not always necessarily benefit PM performance. Similarly, Guynn et al. (1998) found that reminders only improved fulfilment of delayed intentions in some circumstances. More specifically, reminders only improved PM when they referred to both, the target event and the intended action (see Morita, 2006 for a similar result in a different paradigm).

Therefore, by understanding the mechanisms by which we fulfil delayed intentions and the conditions under which they might be beneficial, we can aim to improve individuals' adaptive use of cognitive tools.

Early paradigms investigating cognitive offloading

An earlier study that investigated the effectiveness of setting reminders in fulfilling delayed intentions comes from a diary study designed by Marsh et al. (1998). Their paradigm involved asking participants to come in one day and document all activities they had planned for the coming week. A week later, participants returned and documented the activities that they had actually achieved, and they also provided reasons for any failures to complete their intentions. Using this approach meant that Marsh et al. (1998) could examine multiple intentions that individuals had over the course of a week. Furthermore, the paradigm could address the planning and reprioritization processes that individuals used over an extended period of time. They were also able to examine individuals' use of external memory aids to determine how these might affect completion of delayed intentions.

In their study Marsh et al. (1998) found that participants who changed strategies from using their internal memory resources to instead recording intentions in a daily planner led to more intentions being fulfilled. This finding was supported by Holbrook and Dismukes (2009) who found that individuals who used diary entries to plan intentions were more successful at fulfilling their intentions. Both studies support the notion that offloading intentions into the external environment can improve one's success in fulfilling these intentions. However, both studies contained naturalistic components making it difficult to attribute any differences between offloaded and non-offloaded intentions to the effect of the reminder itself. This is because in real-life there are various reasons for which intentions we might choose to offload. For example, we might be more likely to offload intentions of higher importance than those of lower importance. Therefore, it is difficult to ascertain how diary entries were used to facilitate remembering. Furthermore, it is also difficult to measure how these intentions were planned and reprioritized without these metrics being controlled in an experimental setting.

Therefore, we need to consider how intention offloading can be measured experimentally.

Recent experimental paradigms investigating intention offloading

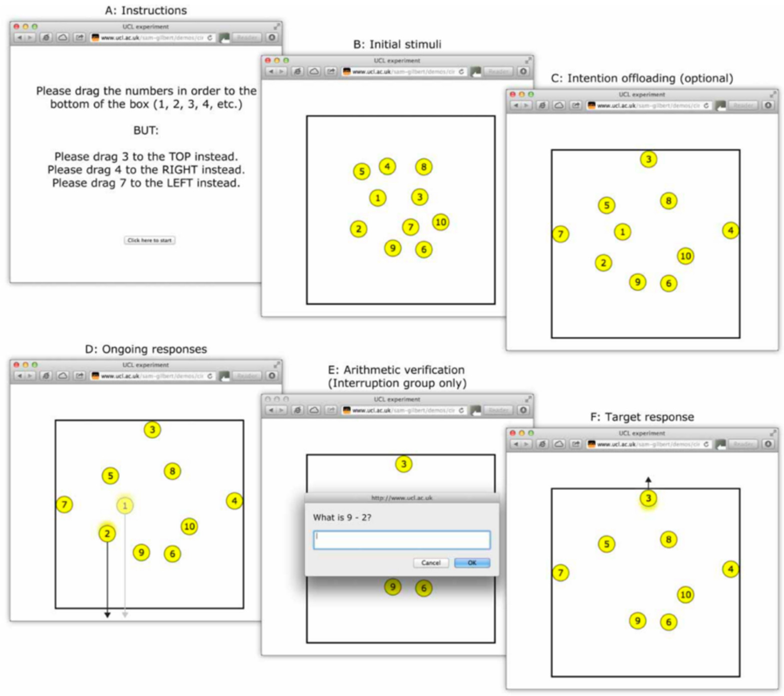
Experimental research has investigated the role of cognitive offloading in a wide variety of domains such as working memory (e.g., Ballard et al., 1995; Grinschgl et al., 2020; Kirsh & Maglio, 1994; Risko & Dunn, 2015), learning (e.g., Hu et al., 2019), visual perception (e.g., Dunn & Risko, 2016) and memory for delayed intentions (e.g., Gilbert, 2015a; Gilbert et al., 2020). Since the next three chapters will have a particular focus on intention offloading which is the memory for delayed intentions, the research detailed below will focus mostly on this domain.

One of the first experimental paradigms investigating intention offloading was developed by Gilbert (2015a). This study was administered as an online web-based task (see Figure 1) where on each trial participants had to perform an ongoing task which involved sequentially dragging 10 numbered yellow circles to the bottom of a square. Alongside this ongoing task, participants were presented with delayed intentions on each trial such that they were instructed to drag either one circle (1-target condition) or three circles (3-target condition) to specific alternative locations within the square (either left, right, or top). This meant that participants had to form delayed intentions that required them to fulfil particular actions when they encountered these predefined cues.

In this paradigm, participants could remember these intentions either by maintaining and rehearsing the intentions internally or by offloading these intentions at the beginning of each trial by dragging the target circles next to their intended locations. Offloading in this

Figure 1

Schematic of the Intention Offloading Task developed by Gilbert (2015a)



Note. Figure adapted from Gilbert (2015a).

way meant that participants no longer needed to maintain this intention internally as the location of the circle provided a perceptual trigger. An everyday analogy of this would be placing an object by the front door so that you remember to take it with you before leaving the house.

The aim of this experiment was twofold. The first was to investigate whether participants decide to use reminders even when they are able to use their own memory. The second was to examine whether intention offloading is influenced by task characteristics which would suggest an influence of metacognitive insight into the likelihood of forgetting. To investigate the latter, the paradigm manipulated two task characteristics. The first was memory load which was manipulated by the number of intentions participants had to remember (either 1-target or 3-targets). The second was task interruption where a pop-up box appeared during the ongoing task asking participants an arithmetic question, so participants had to abruptly pause their response to the ongoing task to input their answer in the pop-up box.

Gilbert (2015a) reported three main findings, 1) participants set more reminders when they had to remember three intentions as opposed to just one, 2) participants set more reminders when they encountered task interruptions, and 3) increased memory demands and greater task interruptions were associated with a decrease in accuracy when participants were forced to use their own memory. The first two findings indicate that intention offloading was influenced by both the memory load posed by the task and the characteristics of the ongoing task in which the intentions were embedded. The third finding suggests that participants' intention offloading strategies were influenced by metacognitive awareness of the likelihood of forgetting. In other words, participants were more likely to set reminders under conditions where forgetting was more likely. Furthermore, individuals who set more reminders fulfilled more delayed intentions. That is, offloading intentions resulted in better task performance.

Biases and optimality in intention offloading

Although the paradigm developed by Gilbert (2015a) allowed for the investigation of how *often* individuals set reminders, it cannot be used to determine how *optimal* individuals are when making these decisions. Setting reminders involves both costs (i.e., the time and effort it takes to set a reminder) and benefits (i.e., increased likelihood of remembering and subsequently fulfilling intentions). For example, it would not make much sense to set reminders for absolutely everything we intend to do. Therefore, we continually make decisions by weighing up the costs versus the benefits.

Two theories have outlined how individuals decide between using internal versus external resources. The first was introduced by Ballard et al. (1992, 1995). Termed *minimal memory*, this approach outlines that individuals tend to be biased towards external resources where possible. This approach was observed using a Pattern Copy Task (previously called the Blocks World Task Ballard et al., 1992, 1995) where it was observed that participants made extensive use of cognitive offloading strategies over internal memorization in a working memory task. This observation led to the conclusion that individuals are systematically biased towards offloading memory processes on to the environment.

The second of these theories was introduced by Gray et al. (2006). Termed the *soft constraints hypothesis*, this approach suggests that instead of having systematic biases towards external resources, individuals tend to choose strategies that maximize their performance in a task while minimizing time constraints. Once again, this was observed using a Pattern Copy Task where participants tended to offload less when temporal costs were high.

However, one of the difficulties in investigating how optimal individuals are when choosing between external versus internal resources is that the costs of using one strategy over the other is not directly comparable (Gilbert et al., 2020). Taking this into account, Gilbert et al. (2020) investigated how individuals choose between internal versus external strategies based on a single metric which was participants' task performance. To do this, Gilbert et al. (2020) adapted the paradigm of Gilbert (2015a) to investigate whether individuals weigh the costs and benefits of using external cognitive tools optimally or whether they show systematic biases towards either strategy (external versus internal).

In this version of the intention offloading task, participants repeatedly chose between using their internal memory processes alone or offloading intentions and improving performance by using external reminders. if participants chose to use their internal memory, they always earned the maximum reward for each target item they remembered. If, however, they chose to use external reminders, they earned a lesser reward for each correctly remembered target item where the value of each target item varied between trials. Therefore, in this paradigm, using reminders incurred both, a cost (i.e., reduced reward for each correctly remembered target item) and a benefit (i.e., increased likelihood of remembering).

In their experiment, Gilbert et al. (2020) found a systematic bias towards using reminders where participants tended to use more reminders than was optimal (also see Ball et al., 2021; Kirk et al., 2021). Furthermore, individual differences in this bias towards using reminders remained stable over time (Gilbert et al., 2020). They also found that participants who had lower accuracy when using their own memory in this task, also tended to set more reminders than those with better ability (Gilbert et al., 2020) suggesting a metacognitive component for a bias towards using reminders.

The role of metacognition in intention offloading

When we form a delayed intention (such as remembering to take a prescribed medication on time), we often need to decide whether to remember this intention using unaided memory or whether to remember this intention by offloading it to external resources. How do we decide which strategy to use?

One factor that might contribute to deciding between these two strategies is metacognition. Metacognition is our ability to monitor and control our cognitive processes (e.g., Nelson & Narens, 1990). With regards to strategy selection (using internal memory versus external reminders) in intention offloading, an individual's strategy selection might be influenced by their metacognitive beliefs associated with their internal memory (i.e., how confident they are in their memory abilities) and external strategies (i.e., how reliable they think the external storage is) (Risko & Gilbert, 2016). Although research has assumed that metacognition is an important factor when it comes to using reminders to support memory for future intentions (e.g., Knight et al., 2005), the first empirical support for this assumption comes from the results of Gilbert (2015a) who found that participants are more likely to set reminders in conditions where they believed their performance might be poorer (i.e., in conditions where there is high memory load or where there are more interruptions during the task).

Given that participants in Gilbert (2015a) set more reminders in conditions that decreased memory performance suggests that metacognitive belief of their own memory abilities triggered intention offloading (also see Weis & Wiese, 2020). However, an alternative explanation for this could simply be that individuals endeavour to minimize the amount of effort it takes to perform a task (Kool et al., 2010). Therefore, direct evidence for metacognitive influence on intention offloading would require research to demonstrate that intention offloading is *predicted* by participants' metacognitive beliefs of their memory abilities.

Evidence for this was found by Gilbert (2015b). In this experiment, participants performed the same intention offloading task as in Gilbert (2015a) in two phases. In the first phase participants performed the task using unaided memory (i.e., they were unable to set reminders). In the second phase participants were permitted to set reminders. In this phase their use of reminders was measured. Before and after each phase of the task, participants had to provide a subjective rating of how well they expected to perform the task using their internal memory, or how well they thought they had performed the task using their internal memory. Gilbert (2015b Experiment 1) found that participants' likelihood of setting reminders was predicted by, 1) their objective unaided ability in phase 1 (i.e., how much they *needed* the reminders) and independently, 2) their confidence evaluations in phase 1 (i.e., how much they *thought* they needed reminders). Furthermore, Gilbert (2015b Experiment 1a) found that participants' metacognitive evaluations predicted their likelihood of setting reminders even when these evaluations were not predicted by objective accuracy. This provides clear evidence for the influence of confidence on intention offloading.

Using the same paradigm, Boldt and Gilbert (2019) conducted an experiment where one group of participants was explicitly instructed on how to set reminders while a second group was not given these instructions. This meant that the second group could only set reminders if they invented the strategy themselves. Their study found that the group of participants who were not informed of the strategy, spontaneously invented one (although, they tended to offload to a lesser degree than the group that was explicitly informed of this strategy). Furthermore, they also found that in both conditions, the tendency to offload was predicted by participants' confidence where those with lower confidence were more likely to set reminders (Boldt & Gilbert, 2019).

Further evidence for the role of metacognition in intention offloading comes from studies using the paradigm developed Gilbert et al. (2020) where participants' bias towards an external strategy was predicted by participants' metacognitive beliefs about their memory ability (see Ball et al., 2021; Gilbert et al., 2020; Kirk et al., 2021). In other words, these studies have found that participants who were more underconfident in their memory abilities displayed a higher bias towards using external reminders. However, it is important to note that other factors might also contribute to strategic intention offloading. For example, a preference to avoid cognitive effort could also be a contributing factor in intention offloading behaviour. This notion is explored in a later empirical chapter of this thesis.

The role of individual differences in intention offloading

Research has shown working memory to predict performance on PM tasks (e.g., Ball et al., 2013; Ball & Brewer, 2018; Smith & Bayen, 2005). One reason for this is that working memory and PM tasks require similar monitoring processes, where a working memory task that has high working memory load and also requires performance monitoring leads to decrements in PM (Marsh & Hicks, 1998). So, if an individual has to remember more intentions or if one has a lower working memory ability, it would be logical to increase reminder use. Indeed Gilbert (2015a Experiment 1) found that task characteristics designed to increase working memory load (i.e., increasing the number of targets that need to be remembered and introducing task interruptions) led participants to set more reminders in the task.

Direct evidence for the role of working memory predicting strategic intention offloading comes from a study conducted by Ball et al. (2021) where they examined how offloading influences memory for delayed intentions in participants with high versus low working memory abilities. They also investigated how this cognitive ability, in turn, influences decisions to offload intentions. The premise of this research comes from the notion that individual differences in working memory ability are driven by two components. The first is an attention component where goal-relevant information is maintained. The second is a memory component that retrieves this goal-relevant information from memory (Unsworth et al., 2014). A parallel process to the one above takes place in intention offloading tasks where participants first have to notice the intention that requires a response (i.e., attention) and then they have to remember the location it needs to be dragged to (i.e., memory) (Ball et al., 2021). Trying to co-ordinate these multiple intentions while performing a difficult ongoing task is challenging. So, individuals with lower working memory ability should compensate for this deficit by setting more reminders.

In Ball et al.'s (2021) experiment, participants completed three versions of the delayed intentions task developed by Gilbert et al. (2020). In addition, participants also completed three complex span tasks that were used to measure their working memory ability. Ball et al. (2021) found that, 1) individuals with higher working memory ability also had better unaided memory for delayed intentions (i.e., they had better memory for intentions without using reminders) and, 2) individuals with lower working memory ability chose to set more reminders in the intention offloading task. Together, these findings suggest that individual differences in working memory ability not only predict unaided memory for delayed intentions, but it also predicts how often individuals choose to set reminders. In other words, individuals with lower working memory ability compensate for this by setting more reminders.

The role of strategy preservation in intention offloading

Another factor that contributes to intention offloading behaviour was shown by Scarampi and Gilbert (2020a Experiment 2). In their experiment, they investigated whether participants' previous experience with setting reminders influenced their decision to offload again. Termed "the Einstellung effect", previous research has found that individuals often repeat strategies that they have previously used (e.g., Bilalić et al., 2008; Schillemans et al., 2009). A similar result is also found in offloading literature where individuals who use the internet as a form of cognitive offloading are more likely to rely on this strategy in subsequent tasks (e.g., Storm et al., 2017).

In their study, Storm et al. (2017) found that using Google to answer a set of trivia questions made individuals more likely to use Google again when presented with a new set of relatively easy trivia questions. This finding suggests that once individuals use the internet to access information, they are more likely to make use of this strategy again to access other information. Therefore, the aim of Scarampi and Gilbert (2020a Experiment 2) was to extend this research to investigate whether the use of reminders in an intention offloading task would increase the likelihood of relying on this strategy in subsequently presented trials.

In their task, participants performed a variation of the task developed by Gilbert (2015a) where they completed the task in two phases. In the first phase, participants were forced to either use external reminders or they were forced to use their unaided memory. In the second phase, participants were free to choose whether they wanted to set reminders or use their own memory to complete the task. Scarampi and Gilbert (2020a) found that although participants were given free choice, they tended to rely on the strategy that they had used in phase 1. In other words, if participants were forced to use reminders in phase 1, they also chose to use reminders when given free choice in phase 2. The opposite was true if participants were forced to rely on their own memory in phase 1. This finding suggests that previous experience with an offloading strategy influences intention offloading behaviour.

In Experiment 1 of their study, Scarampi and Gilbert (2020a) also found that, at least in the short-term, previously using an offloading strategy did not influence unaided memory in subsequent trials. This finding is relevant to the debate about the potential long-term consequences of using technology to aid cognition.

Consequences of cognitive offloading

With the advent of recent technologies, cognitive offloading has not only become more pervasive, but it has also become more efficient. With smartphones and computers now connected to internet, individuals can save and retrieve a lot of information every day. The internet has rapidly changed what information is available to us and it has also changed the way we find this information and share it with others. The result of this is a *digital expansion of the mind* (Marsh & Rajaram, 2019). Therefore, it is important to consider whether reliance on technologies in terms of cognitive offloading changes cognition. That is, what are the consequences of using technology to offload our memory?

A common argument is that relying on external resources can impair cognitive ability due to a reduction in practicing skills required to perform tasks unaided (Baldwin et al., 2011). This negative effect of offloading memory onto technology was investigated by Sparrow et al. (2011).

In their study, Sparrow et al. (2011) had participants read and type sentences on a computer. Half of the participants were told that the computer would save what they had typed, while the other half were told that the information would be erased. Participants were then asked to recall the sentences without any of them having access to the information that they had typed. They found that participants were more likely to remember information they thought had been erased than when they thought the information would be saved. In other words, when participants thought that the information was stored on an external source, they didn't feel the need to encode it. Henkel (2014) extended these findings where participants were led on a guided tour of a museum and were asked to either take photographs of some items, or they were asked to only observe other items. Henkel (2014) found that participants recalled less about objects they photographed (thus “saved”) than they did about objects they only observed (but see Soares & Storm, 2018 for a different set of results).

The results of Sparrow et al. (2011) and Henkel (2014) imply that saving information (whether it be on a camera or a computer) can make it more difficult to remember that information as individuals likely don’t feel the need to encode information stored on an external source. However, these findings do not account for the benefits of saving information on external sources (Runge et al., 2019; Storm & Stone, 2015). Storm and Stone (2015) theorised that perhaps the costs of saving serves as an adaptive function where once information is saved, cognitive resources can be allocated towards other information. In other words, once information is saved externally, individuals might be in a better position to remember other pieces of information.

The premise of this lies in research on directed forgetting which has found that telling participants to forget a previously remembered list of words can enhance memory for a second list of items (see Bäuml et al., 2010; Sahakyan et al., 2013 for a review). One explanation for the directed forgetting effect is that a *“forget”* cue reduces proactive interference allowing new information to be better remembered than it would have been otherwise (Sahakyan et al., 2013). Taking this finding from the directed forgetting literature into account, Storm and Stone (2015) investigated whether saving information onto a computer (i.e., a *“save”* cue) may have similar effects to directed forgetting. So, because one expects saved information to be available later, there should be less need to remember that information than when they expect that information to not be saved. Therefore, proactive interference from the previous list should be reduced as one would expect to see this list later.

In their experiment, participants studied one list of words ('list A'). This list A was either saved onto a computer so that it could be looked at later, or it was simply closed without the file being saved (which would mean that it was not available to be looked at later). Participants then studied a second list of words ('list B') before being given a recall test for both lists. Storm and Stone (2015) found that when the contents of list A were saved and restudied before the test, participants were able to remember a higher proportion of words from that file than when it was not saved. Furthermore, it was found that saving a list before studying a new list of words significantly improved recall of the contents of the new information that was studied (i.e., a *saving-memory enhancement* for the subsequently studied list) (Storm & Stone, 2015). This finding was extended by Runge et al. (2019) who found that the benefits of offloading memory onto external sources was not only limited to memory performance, but could also free cognitive resources for subsequent unrelated tasks where saving a list of words improved participants' performance in modular arithmetic problems (also see Dupont et al., 2022 for a similar finding in intention offloading).

Together, the findings of Sparrow et al. (2011) (and also Henkel, 2014) and Storm and Stone (2015) (and also Runge et al., 2019) suggest that when we offload information onto an external source, it may reduce our memory for offloaded information but enhance our memory for subsequently remembered information and/or performance in subsequent cognitive tasks. These findings have important implications for understanding how technology affects our memory for offloaded information. Rather than suggesting that using technology to offload information is either straightforwardly *"good"* or *"bad"* for our memory, research suggests that there is a complex trade-off between costs and benefits in cognitive offloading. However, seeing as this is a relatively new field of research, it is important to understand conditions under which the saving-enhanced memory effect can be seen. Therefore, the thesis will finish with considering the role of technology in memory for offloaded information and how this might affect our memory for subsequently presented items.

Thesis overview

The literature reviewed above has focused on the causes and consequences of cognitive offloading. This is the main area of focus of the current thesis. Particularly, the next three empirical chapters will utilize the paradigm developed by Gilbert et al. (2020) and will be structured to answer the following questions:

1. The first study reported in this thesis will be presented in Chapter 2 and will look at the role of metacognition in intention offloading. In particular, Gilbert et al. (2020 Experiment 2) found that bias towards using reminders was predicted by participants' erroneous metacognitive underconfidence in their memory abilities. To further establish the role of metacognition in intention offloading behaviour, we investigated whether it was possible to influence participants' confidence using metacognitive interventions and if so, whether these interventions would in turn influence bias towards reminders.
2. The second study reported in this thesis (Chapter 3) aims to investigate the effect of effort-minimization in intention offloading. As mentioned in the introduction, intention offloading can also be influenced by other factors. One such factor could be effort-minimization where individuals might strive to minimize the amount of effort required to perform a cognitive task (Kool et al., 2010). To address this question, the paradigm developed by Gilbert et al. (2020 Experiment 2) was adapted to include performance-based rewards where one group of participants received

monetary incentives based on their performance while the other group only received points.

3. The fourth chapter investigates whether the influence of confidence on participants' bias towards reminders (as found by Gilbert et al., 2020 Experiment 2) can also be predicted by domain-general metacognitive signals. The study used to address this question was completed in three phases where participants completed a pair of perceptual tasks (see Gilbert, 2015b Experiment 2) and the intention offloading task developed by Gilbert et al. (2020 Experiment 2).

The fifth chapter investigates the consequences of intention offloading. The study detailed in this chapter drew on findings from the directed forgetting literature and adapts the paradigm of Storm and Stone (2015) to assess the conditions under which a saving-memory enhancement effect can be found. In the first study of this chapter, the list that could be saved (i.e., participants could either save list A, list B, or no saving was permitted) was manipulated. The order of test presentation was also manipulated where half of the participants were presented with a test for the list A items first and the other half were presented with a test for the list B items first. Doing this meant that the benefits of offloading list A versus the benefits of offloading list B could be compared. This also aided in the investigation of whether the benefit of offloading either list would depend on which list was tested first. The second experiment presented in this chapter then investigated whether individuals have a preference towards saving one list over another.

Finally, the sixth chapter will review the results of all experiments, discuss the importance and implications of the findings, and will pose questions for future research.

CHAPTER 2. INFLUENCE OF METACOGNITIVE INTERVENTIONS ON STRATEGIC INTENTION OFFLOADING

In everyday life, we often form intentions for future actions that need to be executed after a delay. We can choose to remember these intentions by either maintaining them using internal memory or instead, we can choose to offload these intentions on to the external environment in the form of reminders. These reminders might include making notes, using smartphones or smart devices, or physically manipulating our environment by strategically placing objects as reminders (see Risko & Gilbert, 2016).

Recent research has begun to investigate how and when people decide to use external devices to support their memory for delayed intentions (Ball et al., 2021; e.g., Gilbert, 2015a, 2015b; Gilbert et al., 2020; Kirk et al., 2021). The aim of this line of research is to understand how individuals choose between relying on their internal memory ability and using external resources to remember delayed intentions. One finding from research in this area is that individuals' confidence in their memory abilities (i.e., metacognition) predicts their decision to use reminders (e.g., Gilbert, 2015b; Kirk et al., 2021).

Leading on from this finding, the experiment presented in this chapter aimed to explore whether it is possible to find metacognitive interventions which influence participants' confidence in their internal memory abilities and whether this, in turn, influences their decision to use external reminders. Such a finding would support the role of metacognition in cognitive offloading. This could lead to the development of interventions that optimize individuals' use of cognitive tools to support cognition.

Metacognition and optimality in cognitive offloading

Risko and Gilbert (2016) proposed a metacognitive model of cognitive offloading where the decision to select between offloading and relying on one's internal memory processes is influenced by metacognitive evaluations of both one's internal memory processes and of the external aid. In support of this model, empirical research has shown that one of the contributing factors to cognitive offloading is a belief that performance would otherwise be poor regardless of objective cognitive ability (see Gilbert, 2015b; Risko & Dunn, 2015; Risko & Gilbert, 2016). For example, when navigating to a friend's house, you might be confident that you remember the way to their house and so you might decide to use your internal memory to help you navigate. If you are not confident that you will remember the way, you might decide to rely on Google maps (or Citymapper) to navigate. Offloading one's cognition on to external resources in this way incurs both costs (i.e., time it takes to look up your friend's address and input it into Google maps) and benefits (i.e., successful navigation without getting lost).

Gilbert et al. (2020) examined this notion by investigating whether individuals optimally weigh the costs and benefits of setting reminders, or whether they demonstrate systematic biases towards either strategy. In this paradigm, participants performed a difficult task in which accuracy is low (approximately 50%) when using internal memory, but close to 100% when using external reminders. Participants were given a series of choices between earning a maximum number of points when they used their own memory (10 points per remembered item), or a smaller number point of points (between 1-9) when they decided to use reminders. This allowed them to examine the optimality of choice behaviour. For example, if a participant can achieve 65% accuracy using their internal memory and 100% accuracy using reminders, it would be optimal for them to choose internal memory when offered 6 points or below per item with reminders and external reminders when offered 7 points or above per item.

In their study, Gilbert et al. (2020 Experiment 1) found that participants were systematically biased towards using reminders even when they could have earned more points using their own memory. This is called *reminder bias*. Additionally, Gilbert et al. (2020) found that individual differences in this reminder bias remained stable over time (Experiment 1) and was correlated with participants' metacognitive bias (which is a measure of their under/overconfidence) (Experiment 2). In other words, Gilbert et al. (2020 Experiment 2) found that participants who were more underconfident in their internal memory ability, tended to display a higher bias towards using reminders (likely due to a belief that their performance would otherwise be poor).

The aim of Gilbert et al. (2020 Experiment 2) was to evaluate whether individuals were intrinsically biased away from using internal cognitive processes or whether they were biased towards external reminders due to a metacognitive miscalibration where they believed that their performance would otherwise be poor. In this study, one group of participants were provided with performance-based rewards while the other group was provided with performance-based rewards and metacognitive advice (i.e., on every trial participants were advised on which strategy would maximise performance). It was found that bias towards using reminders was eliminated in the group that received metacognitive advice but not in the group that was not advised, providing an account for the role of confidence in intention offloading.

However, both groups of participants (the group that was advised and the group that was unadvised) received a financial incentive as opposed to only earning points (as in Gilbert et al., 2020 Experiment 1). This means that participants in the advised group had two factors that might have predisposed them towards choosing optimally. The first was a financial incentive to do so and the second was reduced cognitive demand as they were told which strategy would be more beneficial instead of having to deliberate which strategy to use. Therefore, it is unclear whether providing metacognitive advice was sufficient to eliminate this bias towards using reminders.

Taking this into account, the study reported in this chapter investigated whether influencing participants' confidence in their memory abilities alone was sufficient to remove their bias towards using reminders.

Current study

The study reported in this chapter used the same paradigm developed by Gilbert et al. (2020 Experiment 2) and investigated whether metacognitive interventions designed to influence participants' confidence in their memory abilities would, in turn, influence their bias towards using reminders. If such an effect is found, it would provide strong evidence for the influence of confidence on reminder bias. This study allowed for the examination of whether removing underconfidence alone is enough to eliminate reminder bias (i.e., without the effect of a financial incentive). To investigate these aims, the current study explored two questions: 1) Do interventions that shift confidence also shift reminder bias?; 2) Are these shifts in reminder bias predicted by participants' metacognitive bias?

Two metacognitive interventions were manipulated in a between-subjects design. The first intervention was feedback valence where half of the participants received positive feedback on their performance and the other half received negative feedback on their performance. For this intervention, it was predicted that participants who would receive negative feedback on their performance would be less confident in their ability to perform the task (see Raaijmakers et al., 2017). The second intervention was practice difficulty where half of the participants were presented with an easier version of the task in the practice trials while the other half were presented with a more difficult version of the task. For this intervention, it was predicted that participants who would practice a more difficult version of the task would display higher confidence in their ability to perform the main task due to a metacognitive contrast effect whereby subsequent trials would feel easier (see Pansky & Goldsmith, 2014 for a similar result). However, an opposite effect could also occur where performing a more difficult version of the task might lead to a carryover of lower confidence where participants might display lower confidence on subsequent trials. Regardless of which way the effect goes, the main theoretical prediction for the first question remained the same: any intervention that reduces confidence should increase pro-reminder bias, and vice versa. So, if for example, difficult practice trials increased participants' confidence, these participants would also be less biased towards using reminders. For the second question, it was hypothesised that individual differences in reminder bias would be predicted by participants' metacognitive bias where participants who are more underconfident in their memory abilities would also be more biased towards using reminders.

Before commencing data collection, the hypotheses, exclusion criteria, experimental procedure, and data analysis plan were pre-registered (osf.io/y3n8t). This experiment was published in Gilbert et al. (2020) as experiment 3.

Method

Participants

A total of 268 participants (67 in each of the four experimental groups; *mean age* = 37 years; *SD age* = 11 years; *range* = 21-70; 152 male; 115 female; 1 other) were recruited through Amazon Mechanical Turk (MTurk) (https://www.mturk.com) an online platform in which participants receive payment for their completion of web tasks. Participation was restricted to volunteers aged 18 years or above with a minimum of 90% Mechanical Turk approval rate. It was also restricted to participants who specified their location as USA. This was done to reduce variability within our sample. Ethical approval for this study was granted by UCL Research Ethics Committee (1584/003).

A power analysis was performed using G*Power 3.1 (Faul et al., 2007). To determine the sample size, two studies that corresponded with each of the two metacognitive interventions were selected. The first study chosen for the power analysis was conducted by Pansky and Goldsmith (2014). This study investigated the effects of initial task difficulty in a general knowledge task on participants' confidence in subsequent answers to target questions. The reported effect size (η_p^2) for the influence of initial task difficulty on subsequent confidence was .15 in experiment 1 (significant) and .08 in experiment 2 (not significant). Despite not reaching the conventional threshold for significance, (p = .12), the power calculation for this experiment was conservatively based on the smaller effect size, which generated a required sample size of 47 participants in each group.

The second study chosen for the power analysis was conducted by Raaijmakers et al. (Raaijmakers et al., 2017 Experiment 2). This study investigated whether feedback valence influenced participants' judgements of how much effort they invested in the task. To achieve 80% power to replicate the smallest effect (d = .49) in this study (two-tailed, α = .05), a sample of 67 participants in each group would be required (G*Power). Given that the two interventions in this study were manipulated in a 2x2 between-subjects design, to aim for sufficient power to detect each effect at both levels of the other factor, the study was powered to replicate the smallest effect of d = .49 from Raaijmakers et al. (2017 Experiment

2) in a pairwise comparison between two cells of the 2x2 design. This required 67 participants in each of four groups, translating to a total of 268 participants.

If participants were excluded due to our pre-registered criteria (*n* = 47) (see below), additional participants were recruited so that the final sample consisted of 268 participants with 67 participants in each group. Participation took approximately 60 minutes and participants were paid USD 7.50 as compensation.

Exclusion criteria

Participants were excluded if they satisfied any of the following criteria: 1) accuracy in the forced internal condition was equal to or greater than their accuracy in the forced external condition as this would imply that reminders did not improve performance, making data uninterpretable (*n* = 17); 2) accuracy in the forced internal condition and forced external condition was lower than 10% (*n* = 5) or 70% (*n* = 12), respectively; 3) negative correlation between target value and likelihood of choosing reminders, suggesting random or counter-rational strategy selection behaviour (*n* = 6); 4) reminder bias (see Measures section) score of more than 2.5 standard deviations from the cell mean (*n* = 5); 5) metacognitive bias (see Measures section) score of more than 2.5 standard deviations from the cell mean (*n* = 2).

Design

The current study adapted the paradigm of Gilbert et al. (2020 Experiment 2) to investigate whether two metacognitive interventions, feedback valence and practice task difficulty, shifted metacognitive bias and consequently, reminder bias. Both these factors were manipulated between-subjects in a 2x2 between-subjects design where practice difficulty (easy, difficult) and feedback valence (positive, negative) were crossed to yield four groups (i.e., easy-positive, easy-negative, difficult-positive, difficult-negative).

For the feedback valence intervention, half of the participants were presented with positive feedback on their performance while the other half were presented with negative feedback. Feedback was always veridical but was presented with either positive or negative valence (see Table 1). For the practice difficulty intervention, half of the participants were presented with easy practice trials and the other half were presented with a more difficult set of practice trials (details of this are in the Procedure section).

Procedure

Participants performed a variation of the task used by Gilbert et al. (2020 Experiment 2) where they had to drag a total of 25 yellow numbered circles to the bottom of a square (see Figure 2 for a schematic representation of the task). On each trial, participants were initially presented with six yellow numbered circles randomly positioned within a square. Using their mouse, participants had to sequentially drag the circles to the bottom of the square in numerical order. Each time a circle was dragged to the bottom of the square, a new circle appeared in its original location, continuing the numerical sequence. This continued until all 25 circles were dragged out of the square.

Occasionally, new circles (described as special circles to the participant) initially appeared in blue, orange or pink rather than yellow. These colours corresponded with the left, top and right side of the square respectively. Two seconds after appearing on the screen, their colour faded to yellow matching the other circles. When a special circle appeared (e.g., in blue), it represented an instruction to the participant that it should eventually be dragged to its corresponding side of the square (e.g., to the left) when it was reached in the numerical sequence. So for example, a participant drags 1 to the bottom of the screen where it disappears. An orange 7 appears in its place, fading to yellow after 2 seconds. Meanwhile, the participant drags circles 2-6 to the bottom of the screen before

Table 1

Post-Trial Feedback

Accuracy	Feedback (positive condition)	Feedback (negative condition)
0%	You did not get any special circles correct this time.	Room for improvement. You got all of the special circles wrong.
Above 0%, below 50%	Well done – good work! You are responding well to the special circles.	Room for improvement. You got most of the special circles wrong.
Above 50%, below 100%	Well done – excellent work! You responded correctly to most of the special circles.	Room for improvement. You got some of the special circles wrong.
100%	Well done – perfect! You responded correctly to all of the special circles.	You did not get any of the special circles wrong this time.

Note. Table detailing the post-trial feedback participants received depending on their performance and the feedback group that they were randomised into.

dragging 7 to the top. In this way, a special circle instructed participants to form a delayed intention to drag that circle to a nonstandard location when it was eventually reached in the sequence.

A demonstration of the full experiment can be accessed via the following weblink: https://www.ucl.ac.uk/sam-gilbert/CS1/Demo/WebTasks.html.

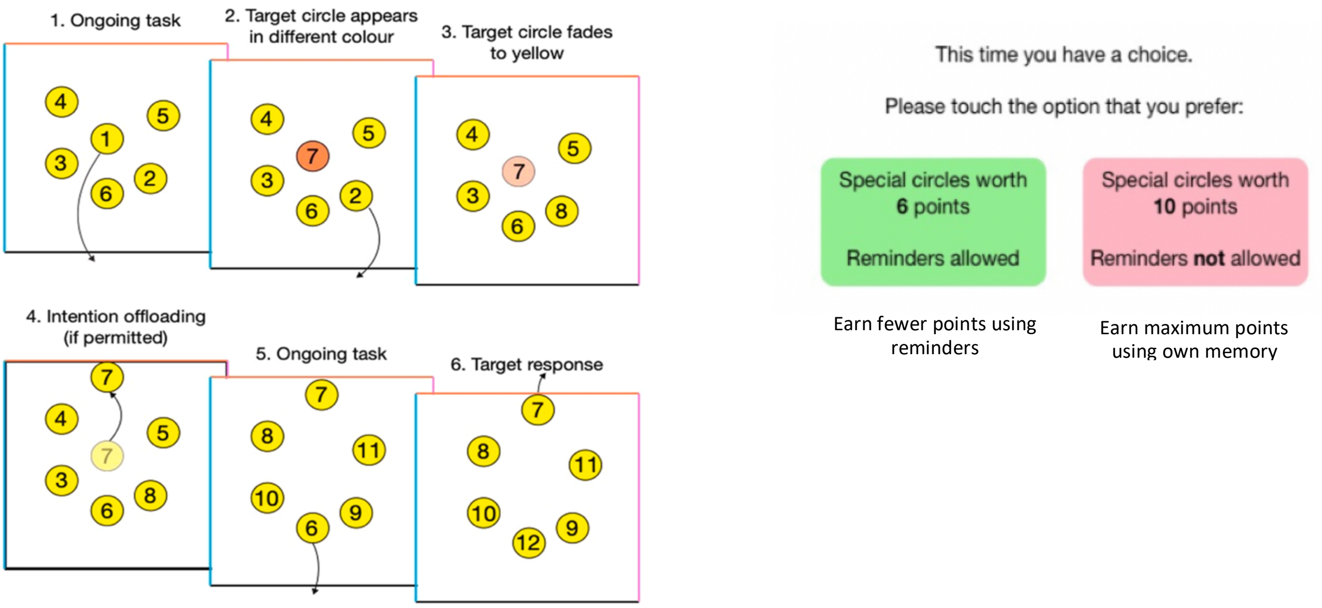
Practice trials

The experiment began with two practice trials. In these two trials, the sequence involved only 7 non-target circles so that participants could practice dragging circles to the bottom of the screen. After completing these two trials, participants performed two additional practice trials involving 8 circles where one of these circles was a target (i.e., special) circle. Participants were instructed on how to respond to this target circle and were only able to proceed past this point if they responded correctly to it (i.e., drag it to its corresponding side

Figure 2

Schematic of the Intention Offloading Task

A) B)



Note. Diagram depicting A) the intention offloading task, and B) the choice trials. Both figures adapted from Gilbert et al. (2020).

of the square) on the second of these practice trials. If they did not, another practice trial was presented.

Following this there were four additional practice trials, each with a sequence of 25 circles. These trials differed between participants based on the practice task condition that participants were randomly assigned to. For participants in the easy-practice trials condition, 4 out of the 25 circles were targets. For participants in the difficult-practice trials condition, 16 out of the 25 circles were targets. Thus, participants in the difficult condition had to remember more target intentions than participants in the easy condition.

In both groups, target circles occurred between circle numbers 7 and 25 in the sequence. These were distributed in a manner that maximized (and equalized) the gap between target circles as much as possible. After these 4 practice trials, and all subsequent trials, participants received post-trial feedback on their performance on that trial. This feedback depended on which feedback valence condition they were randomly assigned to (see Table 1 for details on how feedback was phrased).

After completing these 4 practice trials, participants in the difficult condition were told, “Now the task will get easier”, and participants in the easy condition were told, “Now the task will get more difficult”. They were also told, “It will stay like this for the rest of the experiment. Please ignore the difficulty of the practice trials you have just done and remember that the task will be like this from now on”.

They then received one final practice trial with a total of 10 target circles.

Metacognitive judgement rating

After the practice trial with 10 target circles, participants were asked to provide a confidence rating in their ability to perform the task unaided. They were given the following instructions, “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 it is the same difficulty as the trial you have just completed. The difficulty will stay the same as this for the rest of the experiment. 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 were then presented with a moveable slider on the screen which allowed them to select any percentage between 0-100%.

Intention offloading practice

After participants provided their confidence rating, they were introduced to the offloading strategy. Following this, they were given an additional practice trial where they could practice the offloading strategy.

To move forward with the task, participants had to score higher than 80% (corresponding to 8/10 successfully remembered target circles), otherwise they were asked to repeat the trial. This was done to ensure that participants were able to achieve a high level of accuracy using this strategy.

Scoring points

Participants were then introduced to the procedure for scoring points. They were told that from now on they would score points every time they dragged one of the target circles to its correct location and that they should try to score as many points as possible. After this, they were presented with three practice trials, one for the forced internal trial, one for the forced external trial and one for the choice trial. Unlike Gilbert et al.'s (2020 Experiment 2), the number of points participants earned was not associated with a monetary reward.

In the forced internal practice trial, participants first saw a red button that informed them that every special circle was worth 10 points and that reminders were not allowed. They were also shown the following instruction: "Sometimes when you do the task, you will have to do it without setting any reminders. When this happens, you will score 10 points for every special circle you remember. You will always be given clear instructions as to what you should do. In this case you will be told, 'This time you must do the task without setting any reminders' and you will see a red button. When this happens, the computer will not let you set any reminders. Let's practise that now". After pressing the red button, participants were presented with one practice trial where reminders were not allowed. To force an internal strategy, apart from the next circle in the numerical sequence, all circles were fixed in position on the screen so that participants could not move the target circles when they first appeared.

Participants then received instructions for the forced external practice trial where they saw a green button informing them that every special circle was worth 10 points and that reminders were allowed. They were also shown the following instruction: “Other times, you will have to set reminders for all the special circles. When this happens, you will also score 10 points for every special circle you remember. In this case, you will be told ‘This time you must set a reminder for every special circle’ and you will see a green button. When this happens, the computer will make sure that you always set a reminder for every circle and it will not let you continue if you do not”. After pressing the green button, participants were presented with one practice trial where reminders were allowed. To force an external strategy, participants could only continue the sequence once they moved the target circle within the square.

Please note that while the description above implies that the red button was always associated with internal memory and the green button was always associated with the external memory, this was not the case. In fact, the red and green buttons representing the two strategies were randomized between participants. So for some participants the red and green buttons represented the internal and external strategy, respectively, while for others the association was reversed.

After completing the two forced practice trials, they were asked to practice a choice trial. This was introduced with the following instructions, “Sometimes, you will have a choice between two options when you do the task. One option will be to do the task without being able to set any reminders. If you choose this option, you will always score 10 points for every special circle you remember. The other option will be to do the task with reminders, but in this case each special circle will be worth fewer points. For example, you might be told that if you want to use reminders, each special circle will be worth only 5 points. You should choose whichever option you think will score you the most points. So, if, for example, you think you will earn more points by setting reminders and scoring 5 points for each special circle, you should choose this option. But if you think you will score more points by just using your own memory and earning 10 points for each special circle you should choose this option instead".

After completing the choice trial, participants were able to move on to the main experimental trials.

Experimental trials

Once participants had been familiarised with how scoring worked in the experiment, they were presented with a series of 17 experimental trials. Each of these trials consisted of a sequence of 25 numbered circles with 10 target circles included. On even-numbered trials (i.e., trial numbers 2, 4, 6, 8, 10, 12, 14, and 16), participants performed either the forced internal or forced external condition, in alternating order (with the starting condition counterbalanced between participants). On the remaining 9 odd-numbered trials, participants performed choice trials where they were able to choose between the internal strategy or the external strategy to complete the trial. All possible target values (between 1-9) for the reminders were presented in randomized order.

Before participants started a choice trial, they were presented with a red and a green rectangular button which allowed them to choose between earning the maximum number of points per target with an internal strategy, or a lesser number of points per target using reminders. These two options were presented side-by-side with the left/right ordering of the internal/external options counterbalanced between participants.

Following each trial, participants were informed on the total number of points they had scored in the experiment so far.

Measures

They key dependent measures in this experiment were:

- **Forced internal accuracy (ACC_{FI})**: mean target accuracy (i.e., proportion of target circles correctly dragged to the instructed location) on forced internal trials.
- **Forced external accuracy (ACC_{FE})**: mean target accuracy on forced external trials.
- **Actual indifference point (AIP)**: estimated point at which participants were *actually* indifferent between the two strategy options (i.e., internal strategy and external strategy). As in Gilbert et al. (2020) this was calculated by fitting a sigmoid curve to the strategy choices (0 = own memory; 1 = reminders) across the 9 target values (1-9), using the R package, 'quickpsy', bounded to the range 1-9.
- **Optimal indifference point (OIP)**: target value offered with reminders at which an unbiased individual should be indifferent between the two options, based on the ACC_{FI} and ACC_{FE}. As in Gilbert et al. (2020) this was calculated as:

$$OIP = (10 \times ACC_{FI}) / ACC_{FE}$$

This equation rests on the notion that once we know a participants' mean accuracy when they use either strategy (internal or external), we can calculate their OIP, which is the value attached to the target circles in the external reminder condition that would lead an unbiased individual to be indifferent between the two strategies. For example, if a participant can correctly respond to an average of 5 out of 10 target circles using an internal strategy, and 10 out of 10 using an external strategy, given a choice between 10 points per target with an internal strategy and 9 points per target with an external strategy, it would be practical to select the external strategy. This is because the expected number of points with the internal strategy (10 points x 5 correct responses = 50) would be less than the number of points they would accumulate using the external strategy (9 points x 10 correct responses

= 90). Now if the same participant was given a choice between scoring 5 points per target using an external strategy and 10 points using an internal strategy, the expected number of points when using either strategy would be identical (50 points). Which means that in this case their OIP would be 5 and at this point the participant should be unbiased between the two strategies. In an optimal individual, the OIP and AIP would be equal. Once we have calculated a participants' OIP, we can compare it to their AIP to examine bias towards one strategy or the other.

To derive the equation for the OIP, mean accuracy on the forced external (ACC_{FE}) and internal (ACC_{FI}) trials is calculated. The expected score on internal trials will always be 10 x (ACC_{FI}) as targets in this condition are always worth 10 points. The OIP is the target value that would lead participants to achieve the same score using reminders as they would be using their memory, so the expected score for external trials is OIP x ACC_{FE}. This gives the equation:

$$\text{OIP} \times ACC_{FE} = 10 \times (ACC_{FI})$$

Rearranging this equation to solve for OIP gives us:

$$\text{OIP} = (10 \times ACC_{FI}) / ACC_{FE}$$

- **Reminder bias:** defined as OIP-AIP. A positive number would indicate bias towards using reminders while a negative number would indicate bias away from using reminders.
- **Confidence judgement:** response made on the metacognitive judgement scale using a slider.

- **Metacognitive bias:** difference between metacognitive judgement and actual accuracy on forced internal trials. A positive number would indicate overconfidence while a negative number would indicate underconfidence.

Results

Influence of metacognitive interventions on accuracy

See Table 2 for a summary of results. All analyses were conducted in accordance with the pre-registered plan, except where clearly stated. First, whether the metacognitive interventions influenced participants' accuracy was investigated using a between-subjects ANOVA with factors Practice Difficulty (easy versus difficult) and Feedback Valence (positive versus negative). There was no significant main effect of Practice Difficulty ($F(1,264) = 2.03$, $p = .16$, $\eta_p^2 = .008$) or Feedback Valence ($F(1,264) = .09$, $p = .77$, $\eta_p^2 < .00$). The interaction was also not significant ($F(1,264) = .008$, $p = .93$, $\eta_p^2 < .001$). So, the metacognitive interventions did not have an effect on participants' accuracy.

Influence of metacognitive interventions on confidence

Next, participants' confidence judgement ratings were investigated using a between-subjects ANOVA with factors Practice Difficulty (easy versus difficult) and Feedback Valence (positive versus negative). A significant main effect of Practice Difficulty ($F(1,264) = 5.27$, $p = .022$, $\eta_p^2 = .02$) and Feedback Valence ($F(1,264) = 5.98$, $p = .015$, $\eta_p^2 = .02$) was found where participants who got easier practice trials and positive feedback displayed higher confidence. There was no significant interaction ($F(1,264) = 1.78$, $p = .18$, $\eta_p^2 = .007$).

Then, participants' metacognitive bias (difference between their confidence rating and their accuracy on forced internal trials) was investigated using a between-subjects ANOVA with factors Practice Difficulty (easy versus difficult) and Feedback Valence (positive

Table 2

	Easy positive	**Easy Negative**	**Difficult Positive**	**Difficult Negative**
Forced external accuracy (%)	96.57 (4.84)	95.90 (5.90)	97.69 (4.09)	96.87 (5.66)
Forced internal accuracy (%)	56.53 (16.18)	58.17 (19.35)	59.22 (16.97)	60.56 (19.56)
Confidence Rating	65.85 (23.24)	52.24 (26.81)	53.06 (29.14)	49.15 (34.89)
Metacognitive Bias	9.32 (27.16)	-5.63 (28.40)	-6.16 (29.52)	-11.41 (36.70)
Reminder Bias	1.18 (2.72)	1.78 (2.42)	1.87 (1.95)	2.87 (3.10)
AIP	4.66 (2.51)	4.25 (2.49)	4.18 (2.58)	3.36 (2.78)
OIP	5.84 (1.58)	6.03 (1.86)	6.05 (1.69)	6.23 (1.89)

Note. Table shows means and standard deviations (in parentheses) of reminder bias, metacognitive bias, AIP and OIP in each of the four conditions.

versus negative). A significant main effect of Practice Difficulty (F(1,264) = 8.04, p = .005, η_p^2 = .03) and Feedback Valence (F(1,264) = 7.27, p = .007, η_p^2 = .03) was found. The interaction was not significant (F(1,264) = 1.68, p = .20, η_p^2 = .006).

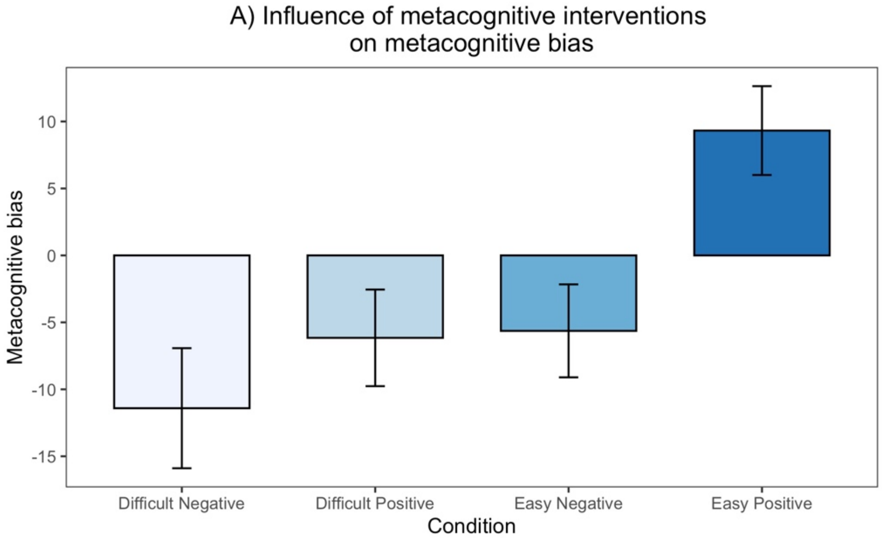
Next, one-sample *t*-tests comparing each group's metacognitive bias score against zero (see Figure 3A) were conducted. It was found that participants in the easy-positive group were significantly (t(66) = 2.8, p < 0.01, d = 0.3) overconfident, while participants in the difficult-negative group were significantly underconfident (t(66) = -2.6, p = 0.01, d = 0.3). Although the easy-negative and difficult-positive groups were underconfident, their means were not significantly different from 0; t(66) = -1.6, p = 0.1, d = 0.2 and t(66) = -1.7, p = 0.09, d = 0.2, respectively.

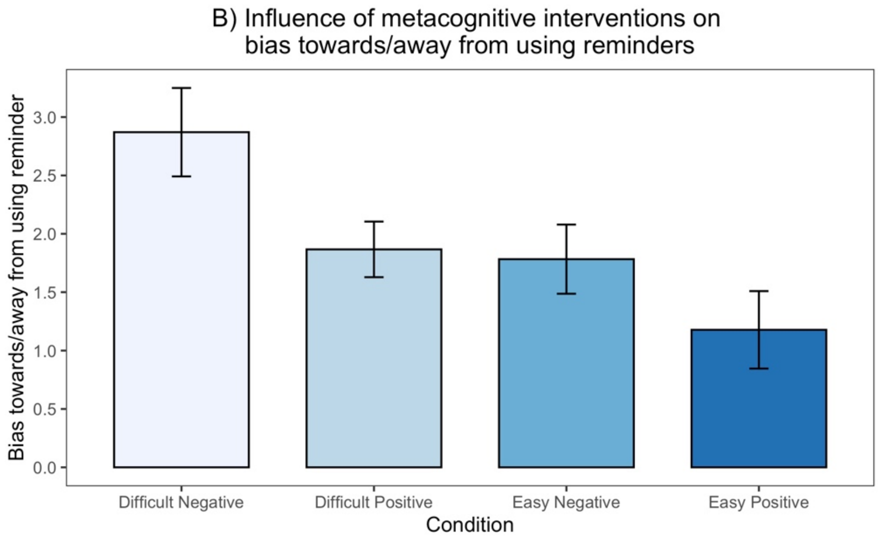
Influence of metacognition and metacognitive interventions on reminder bias

Participants' reminder bias was then investigated in a similar manner to their metacognitive bias. A significant main effect of Practice Difficulty (F(1,264) = 7.93, p = .005,

Figure 3

Figures Depicting Participants Metacognitive Bias and Reminder bias in each of the Four groups





Note. Error bars represent standard errors of the mean. In graph A) positive numbers indicate overconfidence while negative numbers indicate underconfidence. In graph B) positive numbers indicate a bias towards reminders while negative numbers indicate a bias towards internal memory

η_p^2 = .03) and Feedback Valence (F(1,264) = 6.5, p = .01, η_p^2 = .02) was found, but no significant interaction (F(1,264) = .4, p = .53, η_p^2 = .002) was found. One-sample t-tests in all four groups showed that participants were significantly biased towards using reminders (see

Figure 3B) (easy-positive $t(66) = 3.6$, $p < .01$, $d = 0.4$, easy-negative $t(66) = 6.0$, $p < .01$, $d = 0.7$, difficult-positive $t(66) = 7.8$, $p < .01$, $d = 1.0$, and difficult-negative $t(66) = 7.6$, $p < .01$, $d = 0.9$).

To investigate the relationship between metacognitive bias and reminder bias, a multiple linear regression was conducted with reminder bias as the dependent variable and metacognitive bias, practice difficulty, feedback valence and the interaction between practice difficulty and feedback valence as the independent variables. A significant effect of metacognitive bias on reminder bias ($\beta = -.03$, $SE < .01$, $t(263) = -5.8$, $p < 10\text{-}7$) was found. But once metacognitive bias was controlled for in the model, practice difficulty ($\beta = 0.29$, $SE = 0.15$, $t(263) = 1.9$, $p = .054$) and feedback valence ($\beta = 0.26$, $SE = 0.15$, $t(263) = 1.7$, $p = .09$) no longer had a significant effect on reminder bias.

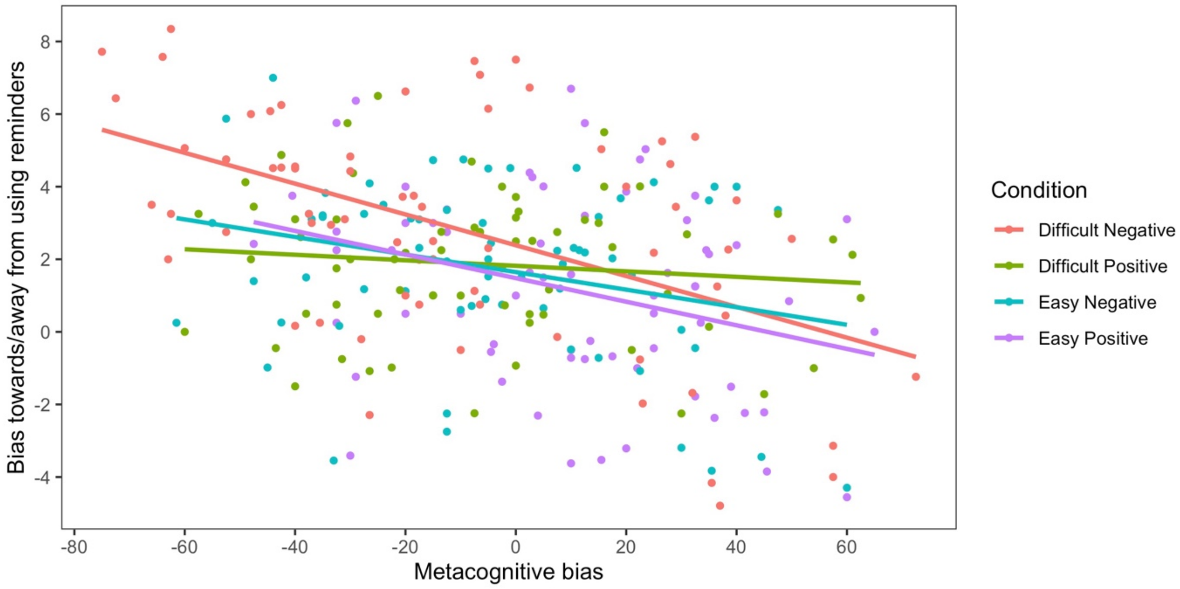
Two mediation analyses using PROCESS (Hayes, 2017) were then conducted on reminder bias to investigate the effects of practice difficulty and feedback valence, with metacognitive bias included as a factor in both models. This analysis was not pre-registered. The models showed a significant indirect effect of practice difficulty ($\beta = .16$, $SE = .06$, $Z = 2.5$, $p = .01$) and feedback valence ($\beta = .15$, $SE = .06$, $Z = 2.4$, $p = .02$) on reminder bias, mediated by metacognitive bias. Since neither intervention had a significant main effect on reminder bias once metacognitive bias was accounted for, it can be suggested that their effects were mediated by participants' metacognitive judgements. The relationship between metacognitive bias and reminder bias is depicted in Figure 4.

Discussion

This chapter examined whether metacognitive interventions designed to influence participants' confidence in their memory abilities would in turn influence their bias towards using reminders.

Figure 4

Correlation between Reminder Bias and Metacognitive Bias



Note. Graph illustrating the correlation between reminder bias and metacognitive bias in each of the four groups.

To investigate this aim, the paradigm of Gilbert et al. (2020) was adapted to include two interventions, practice task difficulty and feedback valence. Both interventions influenced participants' metacognitive bias without influencing their accuracy. In other words, even though participants' confidence differed in the four groups, their accuracy did not. These metacognitive interventions also had a parallel effect on participants' reminder bias where out of the four groups, participants in the difficult-practice/negative-feedback group were the most underconfident and were also the most biased towards using reminders. Further investigation found that these shifts in reminder bias were significantly mediated by participants' metacognitive bias.

As in Gilbert (2015a, 2015b) and Gilbert et al. (2020), a significant negative relationship between metacognitive bias and reminder bias was found where participants who were more underconfident in their memory abilities displayed a higher bias towards reminders. Therefore, the results supported the hypothesis that metacognitive bias would predict participants' bias towards or away from using external reminders. They also supported the prediction that metacognitive interventions designed to influence confidence can shift these reminder biases.

This study also found that confidence alone was not sufficient to account for participants' reminder bias. In three of the four groups, (i.e., easy-practice/negative-feedback, difficult-practice/positive-feedback, and difficult-practice/negative-feedback), participants were underconfident and displayed bias towards using external reminders. However, this was not true for participants in the easy-practice/positive-feedback group. Participants in this group were simultaneously overconfident and biased towards using reminders. This shows that overconfidence in one's memory abilities does not necessarily remove bias towards using reminders seeing as in this experiment overconfidence was still associated with a bias towards using reminders. Therefore, the results of this experiment demonstrate that although confidence contributes to reminder bias, it is not the only factor contributing to this bias.

Other factors such as a preference to avoid effort might also play a role in reminder bias as using one's internal memory is effortful (Ballard et al., 1997b) and so, participants might have chosen to utilize a less effortful strategy. Research has found that rewards give individuals incentives to work harder (e.g., Fröber & Dreisbach, 2014; Krebs et al., 2010). For example, Krebs et al. (2010) found that participants responded faster and more accurately when expecting a greater reward for naming the colour of a stimulus in the Stroop task. Furthermore Fröber & Dreisbach (2014) found that the prospect of performance-contingent reward promotes cognitive stability and proactive control (also see Jimura et al., 2010; Locke & Braver, 2008; Padmala & Pessoa, 2011). Support for this factor also comes from the results of Gilbert et al. (2020 Experiment 2 advised group) where reminder bias was eliminated in the group that received both a financial incentive and metacognitive advice. This notion is further explored in Chapter 4 of this thesis.

With the exception of the easy-practice/positive-feedback group, participants in this experiment were underconfident in their memory abilities, and in all four groups there was a bias towards using reminders. Even though participants in the easy-practice/positive-feedback group were simultaneously overconfident and biased towards using reminders, reminder bias in this group was less than that of other groups. This means that interventions designed to improve metacognitive insight can, to some degree, guide individuals towards more effective use of tools (this notion will be discussed further in the General Discussion section).

With the advent of modern technology, there are numerous opportunities to supplement memory for delayed intentions onto external tools. However, such supplementation can only be beneficial if individuals can correctly judge the optimality of setting reminders in the first place. An example of where over-reliance on technology can sometimes lead to adverse effects is in aviation where Lee and See (2004) reported an incident where pilots trusted the ability of the autopilot feature and failed to intervene by taking manual control when the autopilot function crashed the aircraft that they were flying. Researchers in the field of human factors have termed this “automation bias”, where individuals place excessive trust in the capabilities of automation tools and become complacent (see Parasuraman & Manzey, 2010). Lee and See (2004) hypothesized that trust in external tools and confidence in one’s own abilities are key factors in automation bias.

However, the opposite effect (i.e., under-reliance on technology) can also have negative consequences. For example, Cauvin et al. (2019) found that young adults are generally overconfident outside of a laboratory setting, leading them to inaccurately predict what items they would be able to recall. Another avenue where individuals might display overconfidence is in that of clinical disorders. For example, Knight et al. (2005) found that compared to healthy adults, patients with traumatic brain injuries fail to update metacognitive evaluations of their abilities leading to overconfidence in their PM performance. This overconfidence in one's memory abilities might lead to inadequate use of external aids. Research has found that external reminders can substantially increase ability to remember delayed intentions in people with PM failure (Fish et al., 2010; B. A. Wilson et al., 2001). This emphasises the importance of improving metacognitive insight during rehabilitation (J. Fleming et al., 2017; Vorwerk et al., 2022b). Therefore, an interesting avenue for future research would be to investigate the effectiveness of training individuals' metacognitive insight to optimally utilize compensatory strategies.

Collectively, the findings of this chapter add to the growing body of evidence supporting the role of metacognition in cognitive offloading where confidence has been shown to guide offloading strategies in various domains such as memory for delayed intentions (Boldt & Gilbert, 2019; see Gilbert, 2015a, 2015b; Gilbert et al., 2020; Kirk et al., 2021) and perception (see Dunn & Risko, 2016).

Conclusion

In conclusion, the present results add to our understanding of the role of confidence in cognitive offloading. This study found that metacognitive interventions designed to shift confidence can also influence reminder bias. This reminder bias is, in turn, was mediated by these shifts in individuals' confidence. However, confidence is not the only factor influencing reminder bias. Understanding factors influencing this bias, and interventions that might reduce it can improve individuals' use of external resources as we become increasingly reliant on technology.

CHAPTER 3. THE ROLE OF EFFORT-MINIMIZATION IN COGNITIVE OFFLOADING

In the last chapter we saw that people's confidence in their cognitive ability influences their decision to set reminders (also see Boldt & Gilbert, 2019; Gilbert, 2015b; Gilbert et al., 2020; Kirk et al., 2021). However, the results of the last chapter suggest that confidence cannot exhaustively explain decisions to offload. This was seen in the easy-practice/positive-feedback group where participants were simultaneously overconfident in their cognitive ability but still biased towards using reminders. This shows that metacognition cannot be the only factor influencing decisions to offload.

An alternative reason as to why individuals might decide to offload is to reduce effort associated with remembering intentions internally. It is important to note that while the accounts of confidence and effort minimization might be conceptually different, they might not be mutually exclusive.

The main aim of the current chapter was to investigate the role of confidence and effort minimization in cognitive offloading.

Biases and optimality in cognitive offloading

As discussed in chapter 2, Gilbert et al. (2020) developed an experimental paradigm to investigate whether participants optimally balance the costs and benefits of setting a reminder. In experiment 1 of their study, they found that participants were systematically biased towards using reminders where they tended to choose reminders even when they could have earned more points using internal memory.

In experiment 2 of their study (Gilbert et al., 2020 Experiment 2), a between-subjects design was employed where one group of participants received metacognitive advice (i.e., on each trial they were informed whether it would be optimal to choose an internal strategy or an external strategy based on their performance) while the other group did not. Both groups however, received performance-based monetary rewards. The aims of this experiment were twofold. The first was to investigate whether participants have an intrinsic bias against cognitive effort. In this case providing them with metacognitive advice should not affect their bias towards reminders (i.e., reminder bias). The second was to investigate whether this reminder bias is influenced by one's underconfidence (even when they are provided with a financial incentive to choose optimally). In this case, reminder bias should be eliminated in the group that was provided with metacognitive advice. Gilbert et al. (2020 Experiment 2) found that participants who were not given metacognitive advice displayed a bias towards using reminders even when they had a financial incentive to choose optimally. However, in the group that received metacognitive advice, this bias towards reminders was eliminated thus lending support to the notion that bias towards reminders arises from inaccurate metacognitive evaluations of one's internal cognitive abilities. They also found that participants' reminder bias was influenced by their metacognitive bias (i.e., discrepancy between their confidence ratings and objective performance showing under/overconfidence). In other words, participants who were more underconfident in their cognitive abilities were more biased towards using reminders.

It should be noted (as it was in Gilbert et al., 2020 Experiment 2) that in this experiment it was unclear as to whether simply removing underconfidence would be sufficient to eliminate reminder bias. More specifically, the group of participants who received metacognitive advice had at least two factors that might have led to the elimination of reminder bias. The first factor was a financial incentive to choose optimally. The second factor was reduced cognitive demand as they were informed on which strategy would be optimal based on their performance thus removing the need to decide between the two strategies. Therefore, the elimination of reminder bias in this experiment could simply be explained by a reduction in cognitive demand as opposed to metacognitive error. To differentiate between the two accounts, both would have to be manipulated within a single study.

With regards to the metacognitive account, chapter 2 extended the findings of Gilbert et al. (2020) and investigated whether the effect of two interventions designed to influence metacognitive judgement (i.e., metacognitive interventions) would also influence reminder bias. The results showed that these metacognitive interventions shifted reminder bias in a manner that was mediated by shifts in their confidence thus lending support to the metacognitive account of cognitive offloading. However, this influence of metacognitive interventions on reminder bias where participants displayed underconfidence and a bias towards using reminders was only found in three out of the four groups in this experiment. In the fourth group (easy-practice/positive-feedback), participants simultaneously displayed significant overconfidence in their memory abilities *and* a bias towards using reminders. If confidence is the only factor related to reminder bias, one would expect underconfident individuals to use too many reminders and overconfident individuals to use too few. However, the results of chapter 2 showed that reminder bias can be observed in the context of both under-and-over-confidence.

One possible explanation for this could be that the metacognitive measure used in chapter 2 was incorrect as only one metacognitive judgement scale was presented to participants after the completion of the practice trials. It is possible that as the experiment progressed, and participants completed more trials, they became increasingly underconfident in their performance as the distance from the interventions used was large enough to stop influencing behaviour.

Indeed, West and Mulligan (2019) demonstrated that prospective metamemory (i.e., confidence in one's prospective memory abilities), like retrospective metamemory (i.e., confidence in one's memory of past events or experiences) displays underconfidence with practice so that the more practice one gets, the more underconfident one becomes in their cognitive abilities. This is called the underconfidence with practice effect (UWP) (Koriat et al., 2002). The UWP effect could explain the pattern of biases observed in the easy-practice/positive-feedback group where participants might have been overconfident in the beginning of the experiment, but slowly became underconfident as they gained more practice.

To investigate this possibility, this chapter included a second metacognitive judgement scale at the end of the experiment to examine whether there was a change in participants' confidence from the beginning to the end of the experiment.

Cognitive effort and reward

Another explanation for the findings of chapter 2 is that, in addition to metacognitive confidence, there might be one or more additional factors that contribute to reminder bias. One potential factor could be a preference to avoid cognitive effort (i.e., the intrinsic bias account from Gilbert et al., 2020 Experiment 2).

The concept of cognitive effort has proved to be quite difficult to define (Shenhav et al., 2017). Instead, effortful tasks are typically defined as a task being difficult or demanding, or a task that gives rise to relatively poor performance (i.e., high response times or low accuracy; see Gilbert et al., 2012 for a discussion on the concept of task difficulty). Research has suggested that effort is aversive (Dreisbach & Fischer, 2015; Kurzban, 2016; Saunders et al., 2017; Shenhav et al., 2017) and that individuals tend to avoid effortful tasks (Kool et al., 2010). Called the "law of less work", when given a choice between options that are attributed with similar levels of reward, organisms typically learn to avoid the option that requires more work or effort (Hull, 1943).

This notion is consistent with the view that individuals have an intrinsic drive to avoid using internal memory resources, and to instead rely on using external resources (Ballard et al., 1997b). However, there is evidence suggesting that cognitive effort is not always costly, and in some circumstances can even be rewarding (Eisenberg, 1992; Inzlicht et al., 2018). Consistent with this notion, research has found that individuals can sometimes show a bias towards using internal cognitive resources rather than external resources (Walsh & Anderson, 2009).

Another account of subjective effort suggests that effortful activities involve cognitive processes (such as those associated with working memory) which are both limited in capacity and potentially applicable to a wide range of tasks across domains (Kurzban et al., 2013). Therefore, to redirect these processes towards other tasks, individuals tend to avoid the expenditure of cognitive effort. This account could explain why remembering an intention internally feels more effortful than simply using an external reminder because to the extent that internal memory capacity is occupied by internally maintained intentions, the use of that capacity for other purposes is hindered. But once an external reminder has been set, this hinderance is eliminated and we are able to pursue other activities.

To examine the role of effort minimization in cognitive offloading, the current chapter investigated whether bias towards reminders can be explained by effort-avoidance. This was done by manipulating financial incentives using performance-based rewards. We predicted that a financial incentive would provide participants with more motivation to expend cognitive effort and would, in turn, reduce their bias towards reminders.

Research in different domains has suggested that rewards give individuals incentives to work harder (see Aarts et al., 2010 for evidence in selective attention; see Padmala & Pessoa, 2011 for evidence in task switching). Since incentives motivate individuals to work harder, they are regularly administered to improve cognitive performance (Botvinick & Braver, 2015). For example, Krebs et al. (2010) found that participants responded faster and more accurately when expecting a greater reward for naming the colour of a stimulus in a Stroop task.

The notion that individuals are willing to expend more effort when rewards are available is called motivational vigor (Berridge, 2004; Niv et al., 2006). In support of this idea, research has found that the prospect of performance-based rewards promotes cognitive stability and proactive control (Fröber & Dreisbach, 2014; Jimura et al., 2010; Locke & Braver, 2008; Padmala & Pessoa, 2011). Further support for this notion comes from the experiment conducted by Kool et al. (2010 Experiment 6B) who used a paradigm in which participants repeatedly chose between two visual stimuli. Each switch between the two stimuli was classified as either high demand or low demand. They found that participants constantly gravitated towards the low demand option, but this bias towards the low demand option was reduced when a monetary incentive was linked with the high effort option.

The evidence outlined above suggests that financial incentives increase effort allocated to tasks. So, if one contributor of reminder bias is effort-minimization, financial incentives should reduce this bias. However, it should be noted that individuals might still have an intrinsic bias against cognitive effort that is not completely compensated by the reward offered. Indeed, Westbrook et al. (2013) found that participants would accept a financial penalty to perform a task that is less cognitively demanding. Therefore, the prediction for the current study was that a financial incentive would reduce reminder bias, but it might not eliminate it.

Current study

The current study extended the findings of chapter 2 where the paradigm was adapted to manipulate performance-based rewards. In chapter 2 it was found that out of the four groups, participants in the easy-practice/positive-feedback group were simultaneously overconfident and biased towards using reminders (i.e., in this case the reminder bias could not be explained in terms of metacognitive error). Therefore, only this condition was replicated in the current experiment.

One group of participants received a base payment without any bonus financial incentive (no-reward group; as in the chapter 2) while the other group received financial performance-based rewards in addition to the base payment (reward group). Furthermore, a second metacognitive judgement scale which measured participants' confidence was also included. This scale was presented to participants at the end of the experiment to examine whether there was a change in their confidence from the beginning to the end of the experiment.

Before commencing data collection, the hypotheses, exclusion criteria, experimental procedure, and data analysis plan were registered in a stage 1 registered report (eventually published as Sachdeva & Gilbert, 2020).

Method

Participants

A total of 208 participants (104 in each group) (*mean age* = 38.34 years; *SD age* = 11.32 years; *range* = 20 – 71 years; 121 male; 85 female; 2 other) were recruited through MTurk (https://www.mturk.com). Inclusion criteria for participants were the same as that described in chapter 2 where participation was restricted to volunteers aged 18 years or above with a minimum of 90% Mechanical Turk approval rate. It was also restricted to participants who specified their location as USA to reduce variability in the sample. Ethical approval for this study was granted by UCL Research Ethics Committee (1584/003). Participation took approximately 60 minutes. Payment was decided according to the group they were randomized to (see second 2.5.1 Reward manipulation).

Power calculation

To determine sample size for this study, a power analysis was conducted for each of the three key research questions. Power analyses were conducted using G*Power 3.1 (Faul et al., 2007). The three research questions in this study were:

1. Can the earlier findings from the easy-practice/positive feedback (see chapter 2) where participants were both overconfident in their internal memory abilities and biased towards external reminders be replicated? The aim was to examine this in the no-reward group only (i.e., the group that did not receive performance-based rewards) seeing as this group replicated the earlier procedure. The previous effect sizes (Cohen's *d*) for the one-sample metacognitive bias and reminder bias were .34 and .44 respectively. To achieve 90% power to detect effects of this size (one-tailed one-sample *t*-tests), the experiment required 76 and 46 participants respectively. Seeing as these sample sizes applied to only one group, taking into account both the groups in this study (assuming equal numbers in each group) a total of 152 and 92 participants would be required. To attain 90% power to detect the smaller effect size, a total of 152 participants would be required.
2. Is the reminder bias reduced in the reward group compared to the no-reward group? In the experiment reported in Chapter 2, participants displayed a positive reminder

bias (M = 1.2, SD = 2.7) despite being overconfident. If financial incentives remove any bias against cognitive effort, participants might be expected to have a bias away from using external reminder, i.e., a *negative* reminder bias seeing as they would be overconfident in their unaided memory abilities. However, a financial incentive might not entirely eliminate a bias against cognitive effort (see Westbrook et al., 2013). Therefore, the power calculation for this question was based more conservatively on a scenario where the reminder bias of the reward group is reduced to zero rather than becoming negative. It was also assumed that both would have the same standard deviation of 2.7. This implies a comparison between two groups with means 1.2 (no-reward group) and 0 (reward group) both of which would have a standard deviation of 2.7. This equates to a Cohen's d of .44. To achieve 90% power to detect an effect of this size with a one-tailed, two-sample t-test a total of 180 participants (90 in each group) would be required. A one-tailed hypothesis was used seeing as the hypothesis was directional and we only wanted to test for a difference in this direction.

3. Are participants less confident on the post-task confidence rating (rating given at the end of the experiment) than the first? West and Mulligan (2019) found an underconfidence with practice (UWP) effect in their prospective memory task with an effect size of η^2_p = .15 (Experiment 2, comparison between Blocks 1 and 2). To achieve 90% power to detect an effect of this size, a sample size of 32 would be required under the most conservative assumption that the repeated measures would be uncorrelated. This was planned as a two-tailed test seeing as participants could also become more confident following practice (see Gilbert 2015b), and we wanted to examine any such effect statistically.

These power calculations suggest that a sample size of 180 was sufficient for adequate power to test for the smallest predicted effect. However, ensuring that the study had sufficient power to test the smallest effect alone does not guarantee sufficient power to test all hypotheses together (Francis & Thunell, 2019). Under a conservative assumption that all three analyses described above are independent (i.e., a participant producing data that is consistent with one hypothesis is not more likely to produce data consistent with the others), the post hoc power associated with each of the three tests was multiplied together. With a total sample size of 180, the power to detect all three effects was 84%. Assuming equal numbers of participants in each group, this needed to rise to 208 to achieve 90% power. Therefore, a total of 208 participants with 104 in each group were tested.

Exclusion criteria

The same exclusion criteria used in chapter 2 were also used in this experiment. To reiterate, participants were excluded if they met any of the following criteria: 1) accuracy in the forced internal condition was greater than or equal to the accuracy in the forced external condition (this would imply that reminders do not improve performance making data uninterpretable) (n = 30), 2) accuracy in the forced internal condition is lower than 10% or accuracy in the forced external condition is lower than 70% (n = 9), 3) negative correlation between target value and the likelihood of choosing to use reminders, which would suggest random or counter-rational strategy choice behaviour (n = 13), 4) reminder bias score more than 2.5 standard deviations from their group mean (considered outliers) (n = 5), 5) metacognitive bias score more than 2.5 standard deviations from their group mean (considered outliers). If participants were excluded for any of these reasons (n = 57), additional participants were recruited so that the final sample consisted of 104 participants in each condition (208 in total).

Design

This task was programmed in Java using Google Web Toolkit version 2.8 (http://www.gwtproject.org) and Lienzo graphics toolbox version 2.0 (http://emitrom.com/lienzo), implemented in Eclipse (https://www.eclipse.org).

The experiment followed the procedure of the easy-practice/positive-feedback condition of the experiment presented in chapter 2 with some changes. First, participants were randomly assigned to one of two groups, reward versus no-reward. Participants in the reward group received payment based on the number of points they scored in the task. Participants in the no-reward group received a fixed payment regardless of their performance. This group replicated the procedure of the easy-practice/positive-feedback group of chapter 2. Second, both groups provided an additional confidence judgement at the end of the experiment (post-task confidence judgement).

Procedure

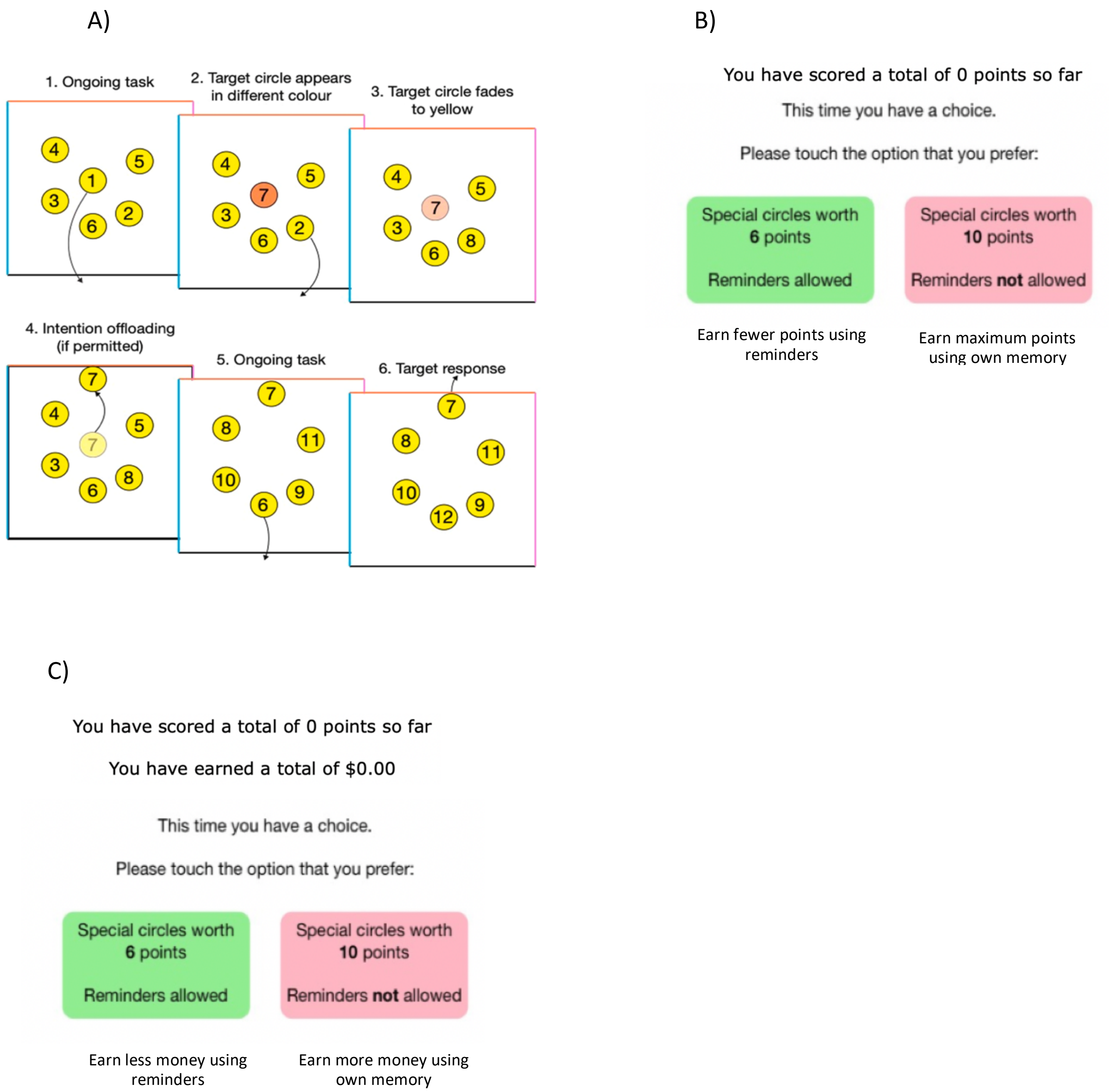
Participants performed a variation of the task used by Gilbert et al. (2020 Experiment 2) where they had to drag yellow numbered circles in sequential order to the bottom a square (see Figure 5 for a schematic representation of the task). The description of the task is detailed in chapter 1 of this thesis (see section Recent experimental paradigms investigating intention offloading).

Reward manipulation

At the beginning of the experiment, participants were randomized into one of two groups: reward or no-reward. Participants in the reward group received the following instruction: “Your payment has not yet been determined. For this experiment, you will earn a base payment of $2.50. Additionally, you will also earn $1 for every 250 points you score. This means that you can earn up to $9.30 for this experiment.” Participants in the no-reward

Figure 5

Schematic of the Intention Offloading Task



Note. Diagram depicting A) the ongoing intention offloading task, B) choice trials for the no-reward group where participants are informed of the total number of points accumulated after each trial, and C) choice trials for the reward group where participants are informed of the total number of points and money accumulated after each trial.

group received the following instruction: “Your payment has now been determined. You will earn a base payment of \$2.50 and an additional \$5 as a bonus for taking part. This means that you will earn a total of \$7.50 for completing this experiment”. The reason for phrasing the conditions in this way was because this experiment was advertised on MTurk as an experiment paying \$2.50 plus bonus. So, participants in the reward group earned a base payment of \$2.50 and an additional bonus dependent on their performance. On the other hand, participants in the no-reward group earned a base payment of \$2.50 and an additional predetermined bonus of \$5 that was not dependent on performance. Based on the results from chapter 2 where participants in the easy-positive group scored an average of 1154 points, this was expected to equate to a performance-dependent bonus of \$4.62. Including the base payment of \$2.50, the amount would total \$7.12. Therefore, even though the maximum potential reward was higher in the reward group, the expected mean earnings were comparable between the two groups.

Practice trials

The practice trials presented in this task were the same as those presented in chapter 2 with the following differences outlined below.

Once participants completed the practice trials and made their pre-task confidence judgements, as in chapter 2 they were familiarised with how the scoring worked in the experiment. Here, the instructions for the two groups differed. As in chapter 2, participants in the no-reward group were told that they would score points for every special circle that they correctly dragged to its corresponding location and that they should try to score as many points as possible. Additionally, they were told that, “points earned will not give you more money, but you should try to score as many points as you can”. Participants in the reward group were told, “You will earn a bonus depending on how many points you score. The more points you score, the more money you will earn”.

As in chapter 2 participants in both groups were then introduced to the forced internal, forced external and choice conditions. Additionally, participants in the reward group were also reminded about the bonus payment with the following instruction, “Please bear in mind that the more points you score, the more you will get paid at the end of experiment”.

Experimental trials

Once participants completed the practice trials and were familiarized with how scoring worked, they completed a series of 17 experimental trials. The sequence of these trials was the same as those described in chapter 2 with the following differences.

First, after completing each trial, participants in the no-reward group were presented with the total number of points they had scored since the beginning of the experimental trials while participants in the reward group were presented with the total number of points and money they had earned.

Second, after finishing the 17 experimental trials, participants were presented with the second confidence judgement scale (post-task confidence scale). The post-task confidence judgement scale was presented with the following instruction: “Now that you have had practice with the experiment, if you are presented with more trials, how accurately do you think you will be able to perform the task without any reminders? The difficulty of these trials would stay the same as the ones you have just completed. Please use the scale below to indicate what percentage of the special circles you will be able to correctly drag to the instructed side of the square, on average, 100% would mean that you can always get every single one correct. 0% would mean that you can never get any of them correct".

Independent variables

The key independent variable in this experiment was the reward group, which was manipulated between-subjects.

Dependent variables

The dependent variables in this experiment were the same as those outlined in chapter 2. These were as follows:

1. **Forced internal accuracy (ACC_{FI}):** mean target accuracy (i.e., proportion of targets correctly dragged to the instructed location) on forced internal trials.
2. **Forced external accuracy (ACC_{FE}):** mean target accuracy on forced external trials.
3. **Optimal indifference point (OIP):** target value offered with reminders at which an unbiased individual should be indifferent between the two options based on their ACC_{FI} and ACC_{FE}. As in Gilbert et al. (2020) and chapter 2 this was calculated as: OIP = (10 X ACC_{FI})/ ACC_{FE}.
4. **Actual indifference point (AIP):** the estimated point at which participants were *actually* indifferent between the two strategy options. As in Gilbert et al. (2020) and chapter 2 this was calculated by fitting a sigmoid curve to the strategy choices (0 = own memory; 1 = reminders) across the 9 target values (1-9) using the R package "quickpsy" bounded to the range 1-9.
5. **Reminder bias:** defined as OIP – AIP, which yielded a positive value for a participant biased towards using more reminders than would be optimal.
6. **Pre-task confidence:** response made to the first confidence scale.
7. **Post-task confidence:** response made to the second confidence scale.

8. **Pre-task metacognitive bias:** difference between the pre-task confidence judgement and accuracy on the forced internal trials. A positive number would indicate overconfidence while a negative number would indicate underconfidence.
9. **Post-task metacognitive bias:** difference between the post-task metacognitive confidence and accuracy on the forced internal trials. A positive number would indicate overconfidence while a negative number would indicate underconfidence.

Results

Influence of performance-based rewards on reminder bias

See Table 3 for a summary of results. All analyses were conducted using R (version 4.0.1). First, whether the earlier findings in chapter 2 from the easy-practice/positive feedback group could be replicated was investigated. These were one-tailed analyses and were only conducted on the no-reward group. A one-same *t*-test (compared to zero) found that participants were overconfident when they made their pre-task confidence judgement ($t(103) = 2.43$, $p = .009$, $d = .24$) and were also biased towards using reminders ($t(103) = 9.63$, $p < .001$, $d = .94$). Therefore, the findings of Chapter 2 were replicated where participants displayed overconfidence in their internal memory abilities and were biased towards using reminders. The same analyses were then conducted on the reward group, but this time they were two-tailed seeing as results in either direction would be theoretically informative. It was found that participants in the reward group were also overconfident ($t(103) = 2.05$, $p = .04$, $d = .20$) and biased towards reminders ($t(103) = 6.44$, $p < .001$, $d = .63$).

Next, the reminder bias between participants in the reward and no-reward groups was compared using a one-tailed independent samples *t*-test. This analysis showed that there was a significant decrease in reminder bias in the reward group compared to the no-

Table 3

	No Reward	Reward
Forced external accuracy (%)	97.24 (4.86)	96.73 (5.19)
Forced internal accuracy (%)	61.95 (16.87)	65.38 (20)
Pre-task confidence rating	67.72 (24.01)	70.48 (24.73)
Post-task confidence rating	64.42 (25.42)	67.81 (25.29)
Pre-task metacognitive bias	5.77 (24.26)	5.10 (25.34)
Post-task metacognitive bias	2.48 (24.16)	2.42 (21.99)
AIP	3.86 (2.61)	5.23 (2.72)
OIP	6.37 (1.69)	6.74 (1.96)
Total reminders used	5.67(2.75)	4.27(2.67)
Total Points	1179.63(183.56)	1250.35(222.82)
Forced trial points	636.73(74.57)	648.46(89.32)
Choice trial points	542.90(122.90)	601.88(142.63)

Note. Behavioural results from both groups. Table shows means and standard deviations in parenthesis. OIP = optimal indifference point; AIP = actual indifference point.

reward group ($t(206) = -2.85$, $p = .002$, $d = .40$) (see Figure 6). This suggests that participants' excessive use of reminders can be explained at least in part, by a preference to avoid cognitive effort.

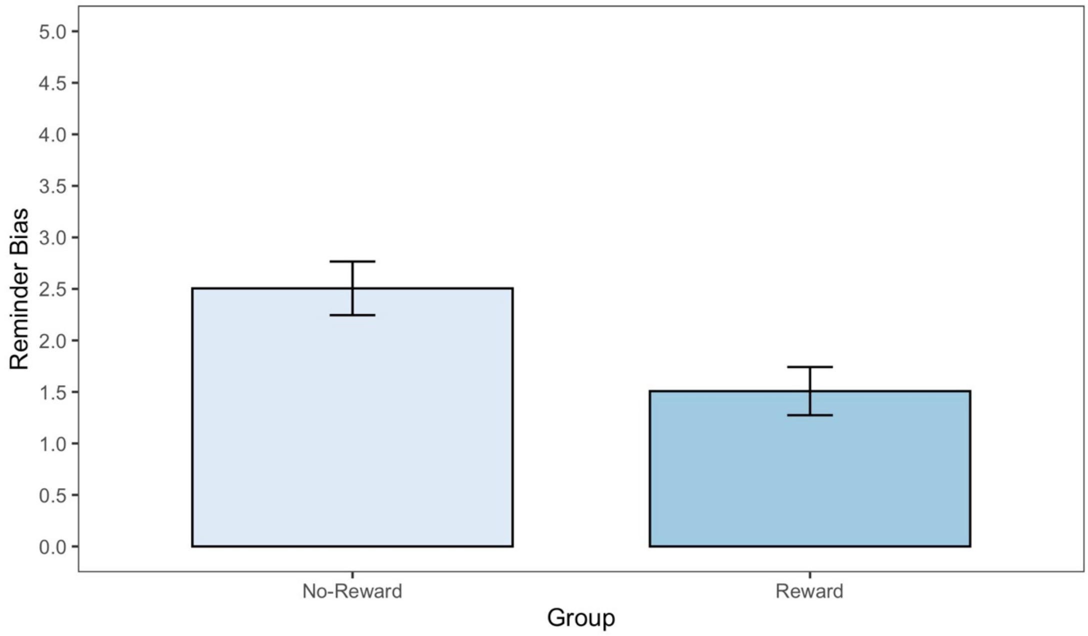
Additionally, the no-reward group also had a higher overall likelihood of choosing reminders over using internal memory. This was reflected in both the total number of reminders used ($t(206) = 3.70$, $p < .001$, $d = .51$) and AIP ($t(206) = 3.71$, $p < .001$, $d = .51$) (see Figure 7A). Also, the total number of points scored (see Table 3) was lower in the no-reward group than in the reward group ($t(206) = 2.50$, $p = .01$, $d = .35$) (see Figure 7B).

Underconfidence with practice effect

To investigate whether participants' confidence changed between the pre-task and post-task ratings, a mixed ANOVA on both confidence ratings using Confidence Judgement Time (pre-task versus post-task) as the repeated measures factor and Group (reward versus

Figure 6

Figure Illustrating Participants' Reminder Bias in each Group



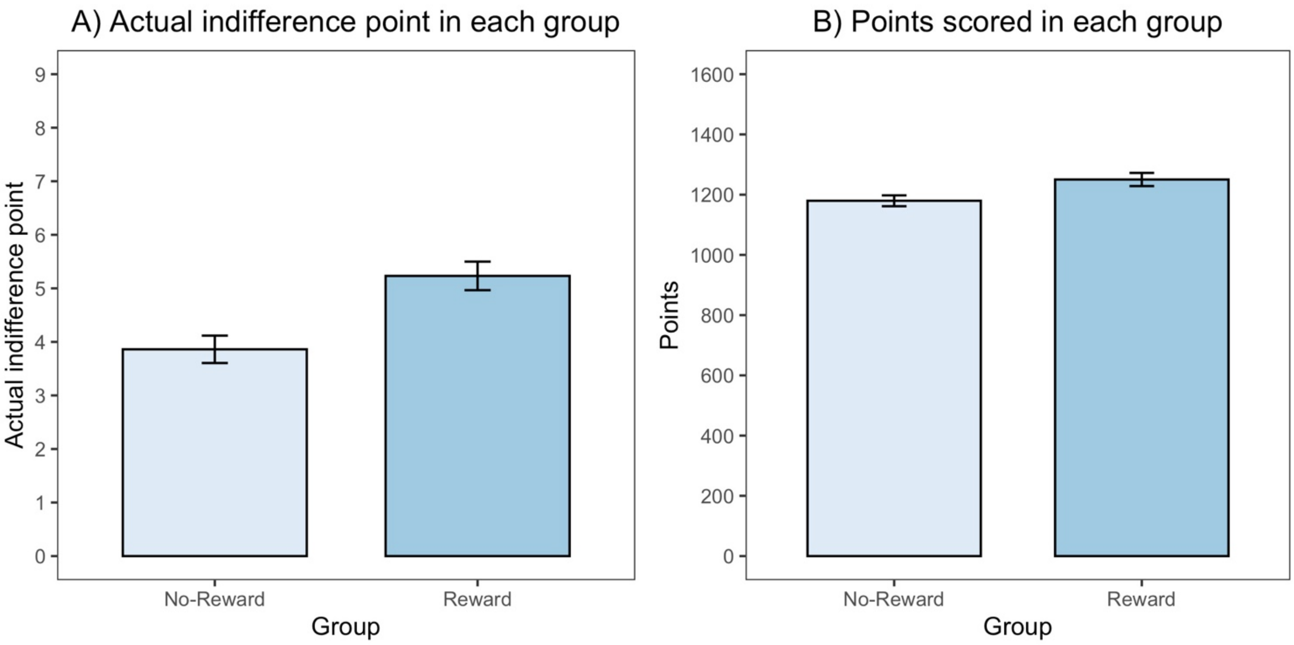
Note. Errors bars represent standard errors of the mean.

no-reward) as the between-subjects factor was conducted. The main effect of Confidence Judgement Time (F(1,206) = 3.43, p = .065, η^2_p = .016) and Group (F(1,206) = 1.02, p = .315, η^2_p = .005) were not significant. The interaction was not significant either (F(1,206) = .04, p = .846, η^2_p < .001). So, the results did not find support for the hypothesis that excessive use of reminders might be caused by a fall in confidence following the pre-task confidence judgement (UWP).

Post-task metacognitive bias with between-subject factor group (reward versus no-reward) was also investigated using a one-way ANOVA. Like the pre-task metacognitive bias, the intercept of this ANOVA indicated overconfidence (M = 2.45), but was not significant (F(1,206) = 2.34, p = .128, η^2_p = .011). The main effect of Group was not significant either (F(1,206) < .0001, p = .987, η^2_p < .0001).

Figure 7

Figures Showing Participants' Actual Indifference Point and Points Scored in each Group



Note. Error bars represent standard errors of the mean. In figure A), a lower actual indifference point represents a lower threshold at which participants decided to set reminders (which means higher use of reminders).

Influence of performance-based rewards on accuracy

Next, whether performance on forced internal and external trials differed between the two reward groups was investigated. A mixed ANOVA was used on target accuracy with Condition (forced internal versus forced external) and Group (reward versus no-reward) as the independent variables. There was a significant main effect of Condition ($F(1,206) = 739.53$, $p < .001$, $\eta^2_p = .782$). But the main effect of Group ($F(1,206) = 1.06$, $p = .31$, $\eta^2_p = .005$) and the interaction were not significant ($F(1,206) = 2.59$, $p = .11$, $\eta^2_p = .01$). So although the groups differed in their strategy choices, their accuracy when using one or the other strategy did not differ.

Relationship between pre-task metacognitive bias and reminder bias

Then, whether reminder bias was related to metacognitive bias was investigated. A multiple linear regression with reminder bias as the dependent variable, and Group (reward = 1 versus no-reward = -1) and pre-task metacognitive bias score as independent variables was conducted. There was a significant main effect of Group ($\beta = -0.50$, $SE = 0.17$, $t(205) = -2.86$, $p < .01$). But the main effect of metacognitive bias was not significant ($\beta = -0.01$, $SE = 0.01$, $t(205) = -1.01$, $p = .31$).

Whether the relationship between reminder bias and metacognitive bias differed between the two groups was also investigated using Pearson's correlations. In both groups, the reward group ($r(102) = -.0076$, $p = .94$) and the no-reward group ($r(102) = -.13$, $p = .18$), metacognitive bias was not significantly correlated with reminder bias. The two correlation coefficients were also compared with each other using Fisher's transformation where both coefficients were transformed to z scores. These correlation coefficients were not significantly different from each other ($z = -.87$, $p = .38$).

Relationship between post-task metacognitive bias and reminder bias

The same analyses were then repeated using the post-task metacognitive bias score instead of the pre-task metacognitive bias score. A multiple linear regression with reminder bias as the dependent variable, and Group (reward = 1 versus no-reward = -1) and post-task metacognitive bias score as the independent variables was conducted. A significant main effect of Group ($\beta = -.50$, $SE = .17$, $t(205) = -2.86$, $p < .01$) but not metacognitive bias ($\beta < -.01$, $SE = .01$, $t(205) = -1.14$, $p = .25$) was found. Furthermore, a Pearson's correlation found that reminder bias was not significantly correlated with post-task metacognitive bias in the no-reward group ($r(102) = -.12$, $p = .23$) and the reward group ($r(102) = -.03$, $p = .75$). Additionally, the correlation coefficients were not significantly different from each other ($z = -.62$, $p = .53$).

Follow-up analysis

In addition to the planned analyses, one follow-up analysis was conducted. The results above suggest that the groups did not differ in their accuracy when they used one strategy over the other, but they differed in their strategy choice. Furthermore, it was found that participants in the reward group earned more points than those in the no-reward group. Since accuracy did not differ between the two groups, the difference in points cannot be attributed to accuracy. Instead, this might suggest that the reward group earned more points because of their strategy choices as opposed to their accuracy when performing the task using one or the other strategy. To test this hypothesis, the total number of points scored on the forced strategy trials and the choice strategy trials were separated. Then, a mixed ANOVA with factors Group (reward versus no-reward) and Trial Type (forced versus choice) was conducted on points scored on the forced and choice strategies. A significant main effect of Group ($F(1,206) = 6.24$, $p = .013$, $\eta^2_p = .029$), a significant main effect of Trial Type ($F(1,206) = 140.12$, $p < .001$, $\eta^2_p = .405$), and a significant interaction ($F(1,206) = 15.87$, $p < .001$, $\eta^2_p = .072$) were found. Further investigation found that when only taking into account the total number of points scored in the forced conditions (forced internal and external combined), the difference between the reward group ($M = 636.73$, $SD = 74.57$) and the no-reward group ($M = 648.46$, $SD = 89.32$) was not significant ($t(206) = 1.03$, $p = .30$, $d = .14$). However, when only taking into account the choice trials, the difference between the reward group ($M = 601.88$, $SD = 142.63$) and the no-reward group ($M = 542.9$, $SD = 122.9$) was significant ($t(206) = 3.19$, $p = .002$, $d = .44$). This suggests that the reward manipulation influenced participants' strategy choice (rather than their accuracy) leading them to score more points.

Discussion

Chapter 2 found that individuals were overconfident in their internal memory abilities but still biased towards using external reminders. This counterintuitive finding suggests that metacognitive confidence cannot exhaustively explain participants' bias towards external reminders. Therefore, in this chapter, individuals' preference to avoid effort was investigated as an additional factor contributing to reminder bias.

In this study, it was hypothesized that providing a financial incentive based on performance would be an additional factor contributing to reminder bias. This hypothesis was supported where reminder bias was reduced in participants who received performance-based rewards.

A secondary aim of this chapter was to investigate whether participants might become increasingly underconfident with practice (Koriat et al., 2002). In other words, this experiment wanted to explore whether participants' initial overconfidence would turn to underconfidence by the end of the experiment. This could also explain a bias towards external reminders. However, this hypothesis was not supported. Although participants were less overconfident at the end of the experiment, their metacognitive ratings did not differ significantly between their initial and their final ratings.

Cognitive effort

Although reminder bias significantly differed between the reward and no-reward groups, both groups chose external reminders more often than optimal. Thus, providing a financial incentive reduced (but did not eliminate) reminder bias suggesting that participants have a bias against cognitive effort that is not fully compensated by financial incentives.

Consistent with this notion, Westbrook et al. (2013) found that participants were willing to accept a financial penalty to perform a task that was less cognitively demanding. Furthermore, research has also shown that the effect of financial incentives on cognitive effort is dependent on the relative value of those incentives (Otto & Vassena, 2020; Rangel & Clithero, 2012; A. Tversky & Simonson, 1993). Therefore, decisions to use internal ability versus external resources might depend on the difference in cognitive effort between the two strategies and the incentives provided. It is also possible that strategy selection is based on one or more additional factors that were not investigated in this experiment.

Further investigating participants' strategy choice, it was found that on forced trials when participants had to use one strategy or the other, there was no significant difference in accuracy between the reward and no-reward groups. However, participants in the reward group earned more points in choice trials likely because they chose to perform the task using unaided memory more often than those in the no-reward group showing that monetary incentives influenced their strategy choice. These results support previous findings where participants are more likely to utilize cognitive effort when there is a financial incentive to do so (Aarts et al., 2010; Botvinick & Braver, 2015; Padmala & Pessoa, 2011). These results also suggest that performance-based rewards influence effort allocation when participants are given a choice of strategy rather than when they are being forced to use one strategy or the other.

Metacognition

One surprising result from this study was that participants' metacognitive evaluations were not significantly correlated with their reminder bias. This contrasts previous findings where such correlations have been repeatedly observed (Boldt & Gilbert, 2019; Engeler & Gilbert, 2020; Gilbert, 2015b; Gilbert et al., 2020; Hu et al., 2019; Kirk et al., 2021). However, it would be premature to draw strong conclusions from this finding given that it rests on a null result and was not the primary aim of the current experiment. One possible interpretation of this result however, could be that the metacognitive interventions reduced the validity of those confidence judgements for predicting individual differences in strategic intention offloading.

Conclusion

Research in cognitive offloading has found that individuals decide whether to use internal cognitive abilities based on various factors such as memory load and task interruption (Gilbert, 2015a; Risko & Dunn, 2015), metacognitive beliefs (Boldt & Gilbert, 2019; Dunn et al., 2016; Dunn & Risko, 2016; Gilbert, 2015b; Gilbert et al., 2020; Hu et al., 2019), participants' past history and previous experience with the act of offloading (Scarampi & Gilbert, 2020b), and objective accuracy (Gilbert, 2015b). The study presented in this chapter found that effort-minimization is another factor that influences decisions to offload as participants were more willing to allocate cognitive effort when presented with monetary rewards. Given that metacognitive evaluations of effort play a role in strategy selection (Dunn et al., 2016), making individuals aware of effort savings associated with cognitive offloading could influence their use of external strategies. This suggests a potential intervention to influence individuals' use of cognitive tools.

CHAPTER 4. DOMAIN-GENERAL METACOGNITIVE PROCESSES IN COGNITIVE OFFLOADING

The results of chapter 2 found a relationship between confidence and intention offloading behaviour where metacognitive interventions designed to influence participants' confidence also (at least in part) influenced their intention offloading behaviour. A relationship between confidence and intention offloading behaviour has also been reported in previous studies where regardless of objective memory ability, participants with lower confidence in their memory abilities are more likely to set reminders in an intention offloading task (Ball et al., 2021; Gilbert, 2015a, 2015b; Gilbert et al., 2020). Similar links between confidence and cognitive offloading in other domains have also been found (e.g., Dunn & Risko, 2016; Hu et al., 2019).

In this chapter, the link between confidence and intention offloading behaviour was further investigated. In particular the extent to which offloading is linked to domain-general versus task-specific confidence signals was examined.

Measures of metacognition

Metacognition has been studied in a wide variety of domains including decision making (e.g., Yeung & Summerfield, 2012), memory (e.g., Nelson & Narens, 1990), strategic intention offloading (e.g., Gilbert, 2015b Experiment 1; Gilbert et al., 2020), and visual perception (e.g., Song et al., 2011). This raises questions of whether metacognitive representations such as estimates of confidence are based on domain-general versus task-specific signals.

Research in metacognition has identified two separate measures of metacognition (S. M. Fleming & Lau, 2014). The first is referred to as *metacognitive bias* which refers to the overall tendency of an individual to report high or low confidence regardless of their performance. The second is called *metacognitive sensitivity* which refers to the ability of an individual to discriminate between different levels of their performance, such as correct and incorrect responses. The current study examined domain-general signals of metacognitive bias in memory and perceptual tasks. Furthermore, metacognitive sensitivity was also measured but only in the perceptual tasks

The domain-general view of metacognition proposes that individuals use a shared metacognitive signal when evaluating their performance across different types of tasks (de Gardelle & Mamassian, 2014; Faivre et al., 2017). In contrast, the domain-specific account of metacognition states that distinct metacognitive resources are leveraged when individuals evaluate their performance across different types of tasks (Morales et al., 2018). By looking at confidence correlations between domains, it is possible to differentiate between the two proposals. For example, if individuals display high confidence in one task and also show high confidence in another task of a different domain, this would lend support to the domain-general account of metacognition (see Baird et al., 2013 for a domain-general account of metacognitive bias). But this is only true if confidence is dissociated from performance (e.g., with a staircase procedure) otherwise correlated confidence might just reflect correlated task performance. If, however, there are no correlations between cross-domain tasks, this would provide support for the domain-specific account of metacognition (see Baird et al., 2013 for a domain-specific account of metacognitive sensitivity).

Understanding the domain-generality of confidence in cognitive offloading could have important practical implications. For example, if cognitive offloading is influenced by domain-general metacognitive signals, this would suggest that a metacognitive intervention (like the ones outlined in Chapter 2) that alters an individual's confidence in one domain, could influence cognitive offloading strategies across multiple domains. In contrast, if cognitive offloading is influenced by task-specific metacognitive signals, this would suggest the need for task-specific metacognitive interventions.

The role of domain-general confidence in intention offloading

Gilbert (2015b Experiment 2) first explored domain-general confidence processes in intention offloading in a web-based task. This study investigated whether, 1) metacognitive bias in two perceptual tasks not only correlated across those tasks but also correlated with confidence in a memory task, and 2) perceptual confidence correlated with participants' likelihood of setting reminders in a memory task (referred to here as offloading proportion).

In this study, participants were presented with a memory task and a pair of perceptual discrimination tasks. On each perceptual trial, participants had to provide metacognitive evaluations of how confident they were that they responded correctly on that trial. A staircase procedure was used to stabilize performance at around 70% accuracy. Using trial-by-trial metacognitive evaluations meant that both measures of metacognition, bias (calculated as mean confidence rating across trials) and sensitivity could be derived. Gilbert (2015b Experiment 2) found that in the two perceptual tasks, confidence and metacognitive sensitivity in one task correlated with its corresponding measure in the other task. This result supports the notion of a shared metacognitive resource between the two perceptual tasks (see de Gardelle & Mamassian, 2014).

This study also found that perceptual confidence correlated with confidence in the memory task. This suggests a domain-general component to confidence where perceptual confidence can predict confidence in a mnemonic task even when it does not predict task performance in the perceptual tasks (as this was equalized using a staircase procedure). Moreover, Gilbert (2015b Experiment 2) found that confidence in perceptual tasks predicted the proportion of reminders participants set in the memory task. In other words, participants who displayed lower confidence ratings in the perceptual tasks also set more reminders in the memory task, which again suggests domain-general signals of confidence across the two task domains.

Metacognitive sensitivity however, was found to be domain-specific as it correlated with its corresponding measure in the perceptual tasks, but did not correlate with any other measure in the perceptual task nor with any of the measures in the memory task (Gilbert, 2015b Experiment 2). Although there are some studies that have found a domain-general component to metacognitive sensitivity (e.g., Mazancieux et al., 2020), the results of Gilbert (2015b Experiment 2) are consistent with the majority of studies in confidence literature, which have concluded in favour of a domain-specific account for metacognitive sensitivity (e.g., Baird et al., 2013, 2015; A. L. F. Lee et al., 2018; McWilliams et al., 2022; Morales et al., 2018).

Current study

The main aim of the current study was to extend the findings of Gilbert (2015b Experiment 2) to investigate whether domain-general confidence signals are linked to individuals' bias towards or away from setting a reminder (i.e., reminder bias).

It should be noted that reminder bias is a different measure to the offloading proportion measure used by Gilbert (2015b Experiment 2). While offloading proportion refers to the proportion of targets that participants set reminders for (Gilbert, 2015a, 2015b) (i.e., their *propensity* to offload), reminder bias refers to bias towards or away from using reminders compared with optimal strategy (i.e., their *preference* to offload). This reminder bias in turn depends on an individual's level of memory performance. So, while a particular reminder setting rate might represent a bias *towards* using reminders in an individual who already remembers well using internal memory, the same reminder-setting rate might represent a bias *away* from using reminders in an individual who performs poorly with internal memory. Although Gilbert (2015b Experiment 2) looked at the relationship between perceptual confidence and the offloading proportion, they did not examine reminder bias.

To investigate whether domain-general confidence signals are linked to reminder bias, the current study employed three tasks. Two of these were perceptual discrimination tasks while the third was a memory task. The memory task was accompanied by two confidence judgement scales. The first one was presented just before the experimental trials (pre-task confidence judgement) and the second one was presented after them (post-task confidence judgement). The perceptual tasks used in this experiment were the same as those in Gilbert (2015b Experiment 2). As in Gilbert (2015b Experiment 2), a staircase procedure was used to stabilize accuracy in the perceptual tasks at around 70%. Using a staircase procedure limits individual differences in task performance which means that individual variation in confidence represents bias rather than true differences in actual task performance.

Instead of calculating metacognitive sensitivity as area under the type II receiver-operating characteristics curve (AUROC2) (as in Gilbert, 2015b Experiment 2), the current study quantified metacognitive sensitivity as metacognitive efficiency using the *M* ratio (meta-*d'/d'*) (Maniscalco & Lau, 2012). The reason for using the *M* ratio was because research has suggested that measures used to quantify metacognitive sensitivity (such as the AUROC2) do not control for the effect of task performance (S. M. Fleming & Lau, 2014). This means that spurious correlations in metacognitive sensitivity might emerge between domains that are driven by variation in task performance rather than metacognitive capacity itself (Rouault et al., 2019).

The *M* ratio attempts to control for this variation in task performance. The *meta-d'* framework models the relationship between performance and metacognition where *meta-d'* is defined as the Type I task performance (*d')* that would lead to the observed type II ROC curve in the absence of noise or imprecision in confidence estimates (Maniscalco & Lau, 2012). *Meta-d'* quantifies the sensitivity of confidence ratings to performance in units of *d'*, which is the signal available for a participant to perform the Type II task (Maniscalco & Lau, 2012). As *d'* and *meta-d'* are quantified in the same units, they can be compared with each other while controlling for task performance.

Hypotheses

The key hypotheses for our study were as follows:

- We predicted a positive correlation between participants' confidence in the perceptual and memory tasks, where participants who predict better performance in the perceptual discrimination tasks would also predict better performance in the memory task.
- We predicted a negative correlation between participants' confidence in the perceptual tasks and reminder bias in the memory task where participants who display lower confidence in the perceptual tasks would also be more biased towards setting reminders in the memory task.
- We hypothesised that reminder bias would have a negative correlation with memory confidence where participants who predict lower confidence in the offloading task would also be more biased towards setting reminders.
- We hypothesised that metacognitive sensitivity would be domain-specific where it would correlate with its corresponding measure across the two perceptual tasks but would not correlate with any measure in the memory task.

Before commencing data collection, the hypotheses, exclusion criteria, experimental procedure, and data analysis plan were pre-registered (https://osf.io/9efjb/).

Method

Participants

A total of 138 participants (88 male; 48 female; 2 other; *mean age* = 25.9 years; *SD age =* 7.7 years; *range* = 18-55) were recruited from Prolific (https://www.prolific.co), an online platform in which participants receive payment for their completion of web tasks. Participation was restricted to volunteers aged 18 years or above. Ethical approval for this study was granted by UCL Research Ethics Committee (1584/003).

To determine sample size, a statistical power analysis was performed using G*Power 3.1 (Faul et al., 2007). The study was powered to detect an effect where the correlation between reminder bias in the intention offloading task could be predicted by confidence in the perceptual discrimination tasks. For this, the result from Kirk et al. (2021) was used where a significant correlation between metamemory bias and reminder bias (r = -.34) was found. Gilbert (2015b Experiment 2) found that offloading was correlated with both perceptual confidence and metamemory. However, the correlation between perceptual confidence and the offloading measure was weaker (r = -.13) than the correlation between metamemory and the offloading measure (r = -.21). So the proportional decrease in strength between perceptual confidence and metamemory from the Gilbert (2015b Experiment 2) study (38%) was used and this decrease was applied to the correlation found by Kirk et al. (2021) (r = -.34). This yielded a new r of .21. To achieve 80% power to obtain an effect of this size (one-tailed test, α = .05), the projected sample size was 138 participants.

If participants were excluded due to the pre-registered criteria (*n* = 11) (see below), additional participants were recruited so that the final sample consisted of 138 participants. Participation took approximately 60 minutes for which participants were paid £7.50.

Design

This task was programmed in Java using Google Web Toolkit version 2.8 (http://www.gwtproject.org) and Lienzo graphics toolbox version 2.9 (http://emitrom.com/lienzo), implemented in Eclipse (https://www.eclipse.org).

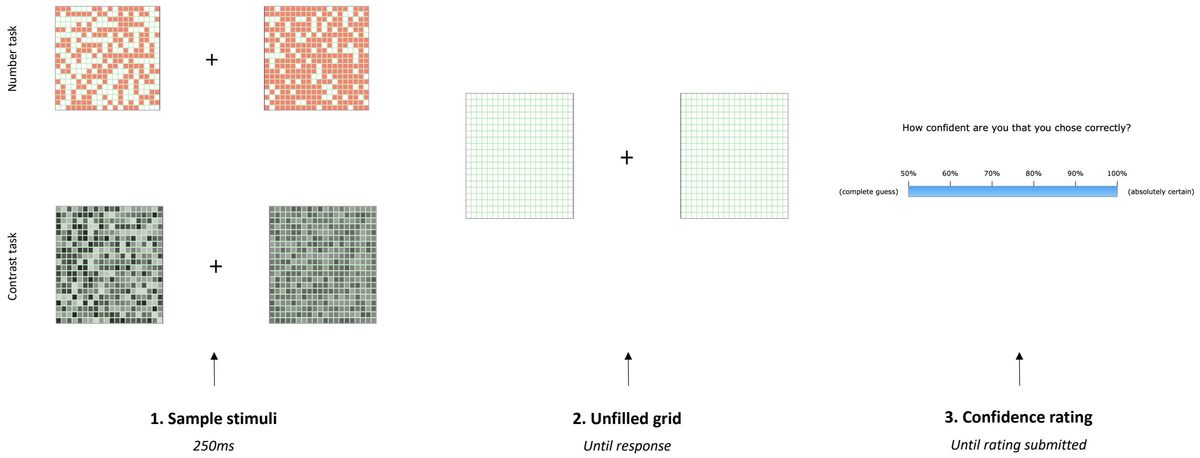
The paradigm used by Gilbert (2015b Experiment 2) and Gilbert et al. (2020 Experiment 2) was adapted to investigate whether confidence can be generalized across memory and perceptual domains.

Participants completed three tasks. One of them was a memory task (see Gilbert et al., 2020; Kirk et al., 2021) and two of them were perceptual discrimination tasks (see Gilbert, 2015b Experiment 2). The order of these tasks was counterbalanced where participants were randomly allocated to perform the memory task before the perceptual discrimination tasks, or vice versa.

Perceptual tasks. Like Gilbert (2015b Experiment 2), the current study had two perceptual tasks, a Number task and a Contrast task (see Figure 8). Accuracy on the perceptual tasks was maintained at about 70% using a two-down-one-up staircase procedure (as in Gilbert, 2015b Experiment 2) where difficulty increased by one step after two consecutive correct responses and decreased by one step if any incorrect responses were made.

In both perceptual tasks participants viewed a fixation point with a pair of grids positioned on either side of it. Both grids were formed of 20 horizontal and 20 vertical pale green lines to yield a total of 400 internal squares. In the Number task, 200 random squares

Figure 8



Note. Schematic representation of the perceptual task

on one side were filled in pink and on the other side more than 200 random squares were filled in pink. Participants were asked to judge which side contained more filled squares. Starting difficulty of this task was set so that 300 squares were filled on one side. To increase the difficulty of the task, this number was gradually reduced so that the side with more squares started to approach 200. Participants received 2 blocks of 75 main experimental trials of this task. A demonstration of the Number task can be accessed here: https://cognitiveoffloading.net/chhavi/numberDemo/WebTasks.html.

In the Contrast task, the squares within each grid were filled with a different shade of grey and participants were asked to judge which grid had a greater contrast between the different shades. Initial difficulty of this task was set so that the shades on one side varied from 15% maximum brightness to 85%, and the shades on the other side varied between 35% and 65%. To increase the difficulty of this task, the difference between the two sides was gradually reduced to approach 25%-75% on both sides. Brightness of the 400 squares was uniformly distributed from brightest to darkest. Participants received 2 blocks of 75 main experimental trials of this task. A demonstration of the Contrast task can be accessed here: https://cognitiveoffloading.net/chhavi/contrastDemo/WebTasks.html.

On each perceptual trial, participants had to discriminate between two perceptual stimuli. After responding, they were asked to give a rating of how confident they were that they responded correctly on that trial. They could make a response by clicking anywhere on a clickable continuous scale. This scale ranged from 50% (their response was a complete guess) to 100% (absolutely certain that they responded correctly). In addition to these trial-by-trial confidence responses, they were also asked to provide an overall post-task confidence rating after completing each perceptual task (see Gilbert, 2015b Experiment 2). This rating was given using a moveable slider ranging from 50% (they thought that every response was a complete guess) to a 100% (they thought that they got every single perceptual discrimination correct) given that chance level for a two alternative forced choice (2AFC) task is 50%.

Memory Task. The memory task used in this experiment was similar to the one in Chapter 2 with two differences. Unlike the previous experiment, feedback was provided in the following form: when a target circle was correctly dragged to its corresponding side of the square box, it turned green before disappearing. When any circle was dragged to an incorrect target location (left, right, or top) it turned red before disappearing. When circles were dragged to the bottom of the square, they always turned purple before disappearing so in this case no feedback was provided. Furthermore, while the previous experiment manipulated the difficulty of the practice trials, in this study practice trials were of the same difficulty as the experimental trials. That is, there were always 10 targets in each trial.

In addition to the memory task, participants were asked for two confidence ratings using a continuous slider ranging from 0% (they thought that they would never respond correctly to any of the target circles) to 100% (they thought that they would respond correctly to all the target circles). One of these sliders was presented to the participants just before the experimental trials (called the pre-task confidence judgement) and the second slider was presented at the end of the memory task (called the post-task confidence judgement). A demonstration of the memory task can be viewed here: https://cognitiveoffloading.net/chhavi/offloadingDemo/WebTasks.html.

Measures

The key dependent measures in this experiment were as follows:

Perceptual tasks

- **Mean accuracy in the Number task:** measure of task performance (i.e., proportion of correct responses) in the Number task.
- **Mean accuracy in the Contrast task:** measure of task performance (i.e., proportion of correct responses) in the Contrast task.
- **Metacognitive efficiency in the Number task:** this measure reflects an individual's metacognitive sensitivity (i.e., how well a participant can discriminate correct from incorrect responses) in the Number task. We used the meta-d'/d' ratio (M-ratio) to calculate metacognitive efficiency (Maniscalco & Lau, 2012). This ratio quantifies metacognitive sensitivity (meta-d') relative to type 1 task performance (d'). Therefore, an optimal value for metacognitive efficiency is 1 as this indicates that all available type 1 information was used in the confidence judgement. To calculate metacognitive efficiency we first discretized the trial-by-trial confidence ratings of each participant into 6 equal bins using quantile ranks and then we fit meta-d' to each participant's confidence rating using the maximum likelihood estimation model

implemented in MATLAB (version 2020a) (http://www.columbia.edu/~bsm2105/type2sdt/) by Maniscalco and Lau (2012).

- **Metacognitive efficiency in the Contrast task:** this measure reflects an individual's metacognitive sensitivity (i.e., how well a participant can discriminate correct from incorrect responses) in the Contrast task. This was calculated using the same method as metacognitive efficiency in the Number task.
- **Confidence in the Number task:** mean confidence across trials in the Number task. This reflects each participants' overall tendency to report high or low confidence irrespective of their performance.
- **Confidence in the Contrast task:** mean confidence across trials in the Contrast task. This reflects each participants' overall tendency to report high or low confidence irrespective of their performance.
- **Post-task confidence rating in the Number task:** additional measure of confidence from a single rating given by participants at the end of the Number task reflecting the proportion of trials on which they thought they were able to correctly discriminate between the two stimuli if they were presented with more trials.
- **Post-task confidence rating in the Contrast task:** additional measure of confidence from a single rating given by participants at the end of the Contrast task reflecting the proportion of trials on which they thought they were able to correctly discriminate between the two stimuli if they were presented with more trials.

Memory tasks

- **Forced internal accuracy (ACC_{FI}):** mean target accuracy (i.e., proportion of target circles correctly dragged to the instructed location) on forced internal trials.

- **Forced external accuracy (ACC_{FE}):** mean target accuracy (i.e., proportion of target circles correctly dragged to the instructed location) on forced external trials.
- **Optimal indifference point (OIP):** target value offered with reminders at which an unbiased individual should be indifferent between the two options. This value was based on the ACC_{FI} and ACC_{FE}. As in Gilbert et al. (2020), the OIP can be calculated as follows: OIP = (10 x ACC_{FI})/ACC_{FE}. In a departure from the pre-registered analysis plan, for each participant the ACC_{FI} measure was derived from just two of the four forced internal trials, randomly selected for each participant. The reason for this was to ensure that independent data were entered into the correlational analyses comparing the OIP with another measure (metacognitive bias) which was also derived from ACC_{FI} (see below).
- **Actual indifference point (AIP):** estimated point at which participants were *actually* indifferent between the two strategy options. As in Gilbert et al. (2020), this was calculated by fitting a sigmoid curve to participants' strategy choices (0 = own memory; 1 = reminders) across the 9 target values (1-9) using the R package 'quickpsy' bounded to range 1-9. The AIP can be seen as an index of offloading frequency (similar to the offloading proportion in Gilbert, 2015b). A low AIP indicates that the participant only required a small target value to choose external reminders and so, they set reminders on a larger number of trials. A high AIP indicates that participant rarely chose external reminders, requiring a high target value to choose to offload.
- **Reminder bias:** defined as OIP-AIP, which would yield a positive number for bias towards using more reminders than would be optimal.

- **Confidence (pre-task):** response made on the pre-task confidence scale. This is participants' tendency to give high or low confidence ratings.
- **Metacognitive bias (pre-task):** difference between the pre-task confidence rating and actual accuracy on forced internal trials. A positive number would indicate overconfidence and a negative number underconfidence. In a departure from the pre-registered analysis plan, this was calculated from the two remaining ACC_{FI} trials that were not used to calculate the OIP, to ensure that the correlation between OIP and metacognitive bias was based on independent data.
- **Confidence (post-task):** response made on the post-task confidence scale. This is participants' tendency to give high or low confidence ratings.
- **Metacognitive bias (post-task):** difference between the post-task confidence rating and actual accuracy on forced internal trials. A positive number would indicate overconfidence and a negative number underconfidence. In a departure from the pre-registered analysis plan, this was calculated from the two remaining ACC_{FI} trials (as above these were randomly selected for each participant) that were not used to calculate the OIP, to ensure that the correlation between OIP and metacognitive bias was based on independent data.

Exclusion criteria

In the memory task, participants were excluded if they satisfied any of the following criteria: accuracy in the forced internal lower than 10%, accuracy in the forced external condition lower than 70%, accuracy in the forced internal trials higher than accuracy in the forced external trials as this would imply that reminders did not improve performance making data uninterpretable (n = 4), a negative correlation between target value and the likelihood of choosing to use reminders, which would suggest random or counter-rational strategy choice behaviour (*n* = 2). Participants were also excluded if their reminder bias exceeded 3 median absolute deviation (MAD) units (C. Leys et al., 2013) (*n* = 1) or if their metacognitive bias exceeded 3 MAD units.

In the perceptual tasks, participants were excluded if their collapsed accuracy in the Number and Contrast discrimination tasks exceeded 3 MAD units as this would suggest a failure of the staircase procedure and/or frequent guessing or random responses (*n* = 4).

Procedure

Before commencing the study, all participants provided informed consent. Once this was provided, they proceeded to begin with either the perceptual discrimination tasks first or the memory task first.

Perceptual tasks. The two perceptual tasks were presented in randomized order. An example stimulus with minimum difficulty was presented for 1000ms and participants were asked to make a response. If an incorrect response was given, another example stimulus was presented. Once participants made a correct response, they were presented with five more trials where stimuli were presented for 800ms. Participants needed to respond correctly to all stimuli otherwise they were asked to repeat these trials. Once five correct responses had been made, participants were presented with five more trials with stimuli presented for 250ms. They needed to respond correctly to at least four of these stimuli before being able to continue the experiment.

Participants were then presented with 40 practice trials and from this point onwards difficulty was adjusted with a two-down-one-up staircase procedure. After these practice trials, participants were introduced to the trial-by-trial confidence scale. Instructions for this were as follows: “Now that you have had some practice with this task, we would like to introduce you to another element. After each discrimination judgement, you will be asked to give a rating of how confident you are that you responded correctly. You will be presented with a scale ranging from 50% to 100%, where 50% means that your response was a complete guess and 100% means that you are absolutely certain that you answered correctly. To give your rating, you can click anywhere on the blue slider". Participants then performed a further 10 trials with the confidence judgement scale.

They then performed 2 blocks of 75 main trials where they were presented with the confidence judgement scale after each perceptual discrimination response. Finally, participants were asked to give their post-task confidence rating with the following instructions: "Now that you have finished this task, we would like you to tell us how accurately you think you can perform the task if there were more trials. Please use the scale below to indicate the percentage of times you can correctly discriminate between the two patterns, on average. 100% would mean that you can always get every single one correct. 50% would mean that every response was a complete guess, like tossing a coin for each answer".

After completing the post-task confidence rating, the other perceptual task (Number or Contrast) was administered in an identical manner.

Memory task. The procedure of the memory task was similar to the procedure of the intention offloading task presented in Chapter 2 with some changes.

Participants were first presented with seven circles without any targets so that they could practice simply dragging the circles to the bottom of the screen. Next, instructions for how to respond to targets was presented and participants performed one practice trial involving eight circles and one target. They were only allowed to continue to the next phase of the task if they responded correctly to the target, otherwise they were asked to repeat the trial.

Following this, participants received two (instead of five) practice trials with 25 circles, 10 of which were targets. After these two practice trials, participants were asked to give their pre-task confidence rating with the following instructions: “Now that you have had some practice with this task, we would like you to tell us how accurately you think you can perform the task. Please use the scale below to indicate the percentage of special circles you can correctly drag to the instructed side of the square, on average. 100% would mean that you can always get every single one correct. 0% would mean that you can never get any of them correct”. After giving their response on the pre-task confidence scale, participants were presented with one practice trial which instructed them how to set reminders by dragging target circles next to their intended side of the square (Gilbert et al., 2020).

After completing the practice trials, participants were introduced to the forced internal, forced external and choice conditions (see Chapter 2). Once participants were familiarized with how the task worked, they were presented with the 17 main experimental trials as described in Chapter 2. After each trial, they were able to see the total number of points they accumulated since the beginning of the main experimental trials.

After finishing the main experimental trials, participants were presented with the post-task confidence scale with the following instructions: “Now that you have finished this task, we would like you to tell us how accurately you think you can perform the task without any reminders if there were more trials. Please use the scale below to indicate the percentage of the special circles you can correctly drag to the instructed side of the square, on average. 100% would mean that you can always get every single one correct. 0% would mean that you can never get any of them correct”.

Once participants completed all three tasks, they were thanked, debriefed, and paid for their time.

Results

We followed our pre-registration plan with one exception. The original pre-registration stated that all statistical tests would be one-tailed, however we realized that this was not described with sufficient clarity, including some cases where the predicted direction of the effect was not clearly specified. Therefore, a conservative approach was taken to report two-tailed tests throughout, with the exception of the correlations between reminder bias in the memory task and confidence in the perceptual tasks. These correlations were clearly specified in the pre-registration as one-tailed tests with a specified direction (see page 3 of the pre-registration), and they were also the basis of the power calculation for determining the sample size. All analyses were conducted using R (version 4.0.3).

Memory task

See Table 4 for a summary of results from the intention offloading task. First participants' metacognitive bias scores were investigated. This was the difference between their responses on the two confidence scales and their accuracy on the forced internal trials. One sample *t*-tests (compared to zero) showed that participants were significantly underconfident when they made their first confidence judgement ($t(137) = 7.61$, $p < .001$, $d = .65$) and when they made their second confidence judgement ($t(137) = 1.99$, $p = .049$, $d = .17$).

Then participants' reminder bias was investigated using a one-sample *t*-test (compared to zero). It was found that participants were significantly biased towards reminders ($t(137) = 12.18$, $p < .001$, $d = 1.04$).

A paired samples *t*-test between the pre-task and post-task metacognitive bias scores was then conducted to investigate whether participants' metacognitive bias changed between the two ratings. Participants (although still underconfident) were significantly less

Table 4

	Mean (Standard deviation)
Forced external accuracy (%)	98.59 (2.43)
Forced internal accuracy (%)	62.34 (16.81)
Confidence rating (pre-task)	44.19 (23.92)
Confidence rating (post-task)	58.86 (23.86)
Metacognitive bias (pre-task)	-18.27 (28.19)
Metacognitive bias (post-task)	-3.61 (21.33)
OIP	6.31 (1.79)
AIP	3.94 (2.46)
Reminder bias	2.37 (2.28)

Note. Table showing means and standard deviations of behavioural results from the offloading task. OIP = optimal indifference point; AIP = actual indifference point.

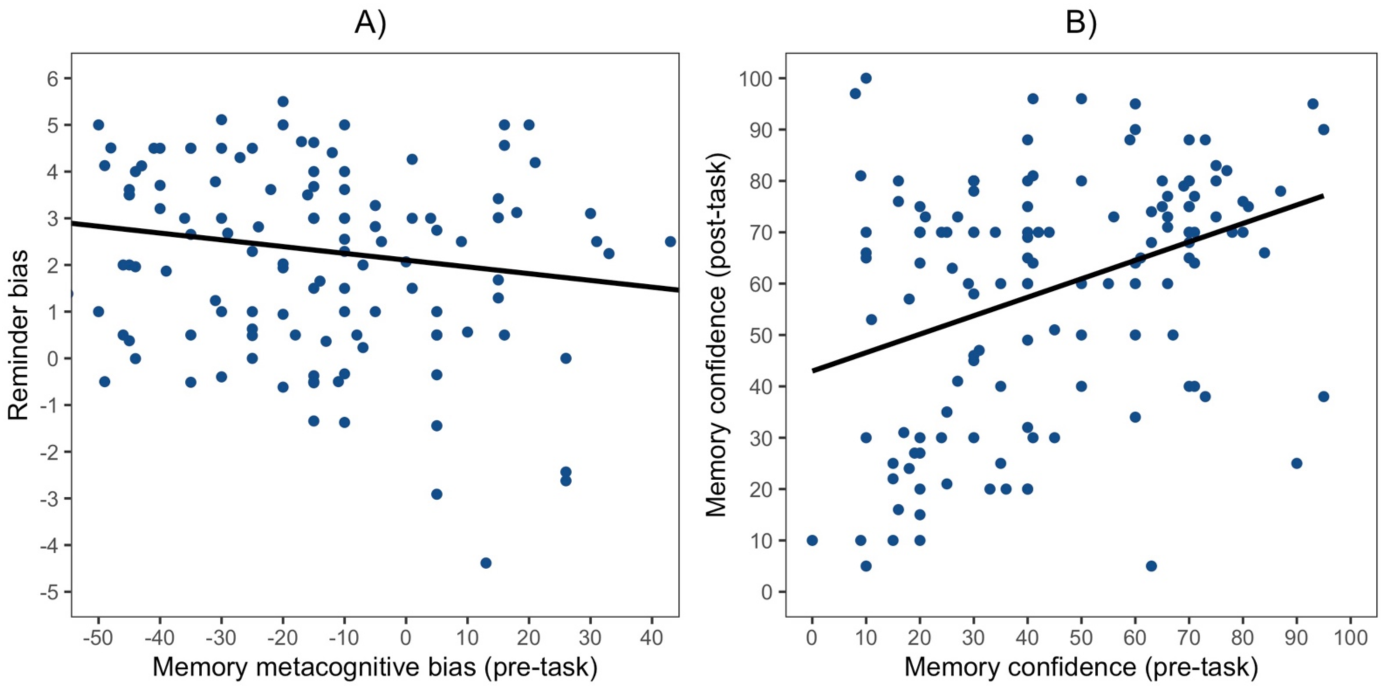
underconfident in their post-task metacognitive bias ratings ($t(137) = 6.37$, $p < .001$, $d = .54$).

To investigate the relationship between reminder bias and pre-task metacognitive bias, a Pearson's correlation was conducted. A significant negative correlation ($r(136) = -.17$, $p = .036$) between these measures was found where participants who were more underconfident in their pre-task confidence ratings displayed a higher bias towards using reminders (see Figure 9A). A similar significant negative correlation was found between participants' reminder bias and their pre-task confidence rating (i.e., their raw confidence rather than under/overconfidence as measured by metacognitive bias) ($r(136) = -.21$, $p = .01$).

Finally, to investigate the relationship between the two metacognitive confidence ratings (pre-task and post-task), a Pearson's correlation was conducted. There was a significant positive relationship between the two confidence ratings ($r(136) = .36$, $p < .001$) where participants who predicted lower confidence in their pre-task rating also predicted lower confidence in their post-task rating (see Figure 9B).

Figure 9

Correlations within the Memory Task



Note. Graphs depicting, A) relationship between reminder bias and pre-task metacognitive bias, and B) relationship between pre-task memory confidence and post-task memory confidence.

Perceptual tasks

See Table 5 for a summary of results. Average accuracy in the perceptual tasks was at around 69.6% indicating that the staircase procedure worked at maintaining participants' accuracy approximately mid-way between chance and ceiling levels.

A paired samples *t*-test showed that accuracy in the Number task was higher than accuracy in the Contrast task ($t(137) = 10.75$, $p < .001$, $d = .91$). However, mean confidence rating ($t(137) = 3.63$, $p < .001$, $d = .31$) and metacognitive efficiency ($t(137) = 2.38$, $p = .02$, $d = .20$) were higher in the Contrast task than in the Number task. Post-task confidence ratings did not differ between the two tasks ($t(137) = .88$, $p = .38$, $d = .07$).

Table 5

	Number task	Contrast task
Accuracy (%)	71.13 (2.2)	68.05 (2.32)
Metacognitive efficiency	0.57 (0.36)	0.67 (0.42)
Confidence	73.09 (9.05	74.82 (9.63)
Confidence judgement (post-task)	69.7 (9.65)	70.46 (11.18)

Note. Table showing the behavioural means and standard deviations (in parentheses) from the perceptual tasks.

Pearson's correlations showed that apart from accuracy ($r(136) = -.10$, $p = .23$), every other measure in the perceptual tasks significantly correlated with their analogous measure in the other task (mean confidence: $r(136) = .82$, $p < .001$; metacognitive efficiency: $r(136) = .27$, $p = .001$; post-task confidence rating: $r(136) = .53$, $p < .001$) (see Figures 10A, 10B and 10C respectively). Furthermore, mean confidence and the single post-task confidence rating at the end of each task were significantly intercorrelated where participants who displayed lower confidence in their mean trial-by-trial ratings also gave lower confidence on the post-task confidence scales ($rs(136) > .45$, $ps < .001$).

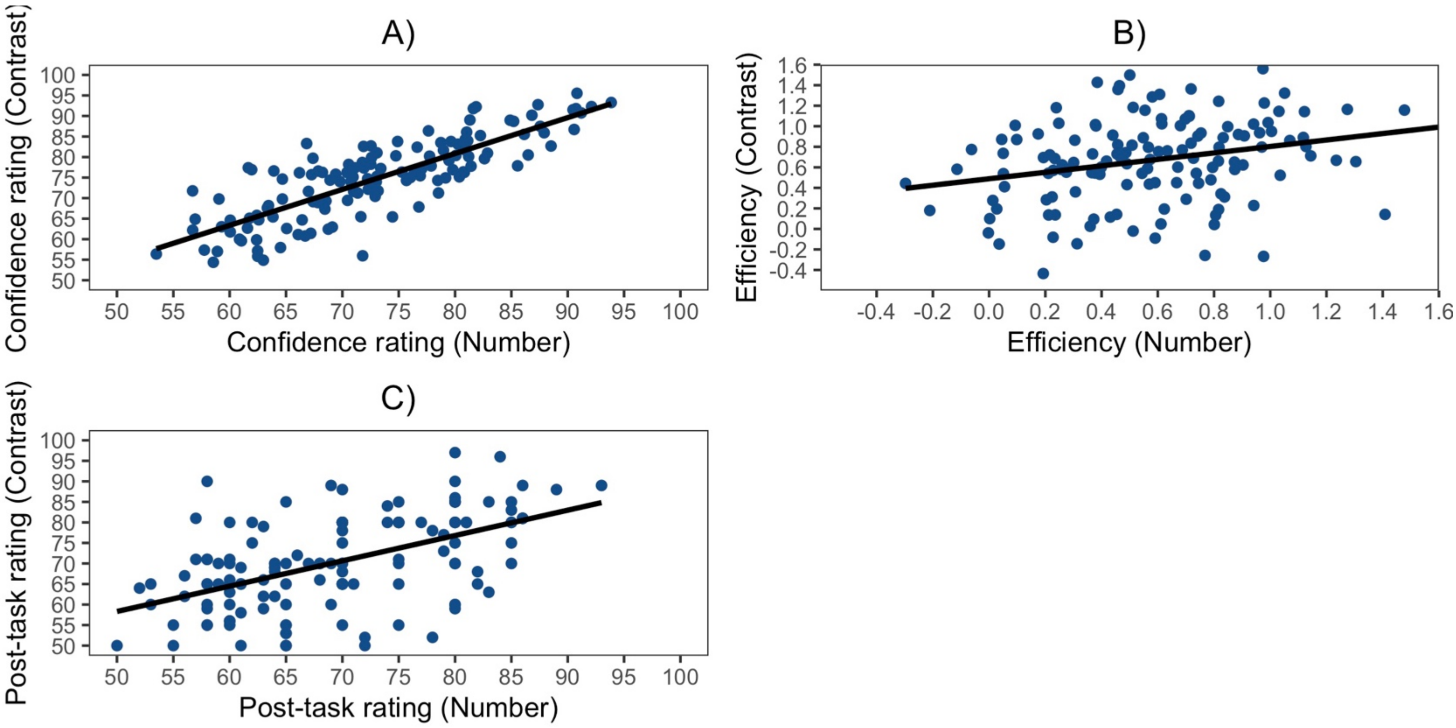
Although metacognitive efficiency in the Contrast task only correlated with its analogous measure in the Number task, metacognitive efficiency in the Number task correlated with mean confidence in both the Number task ($r(136) = -.17$, $p = .04$) and the Contrast task ($r(136) = -.18$, $p = .03$).

Intercorrelations between perceptual and offloading tasks

To investigate the relationship between measures from the perceptual tasks and the memory task, the measures of accuracy, mean confidence, metacognitive efficiency and post-task confidence ratings were collapsed across the two perceptual tasks. To investigate our predictions on the cross-domain metacognitive signals between the perceptual and

Figure 10

Correlations within the Perceptual Tasks

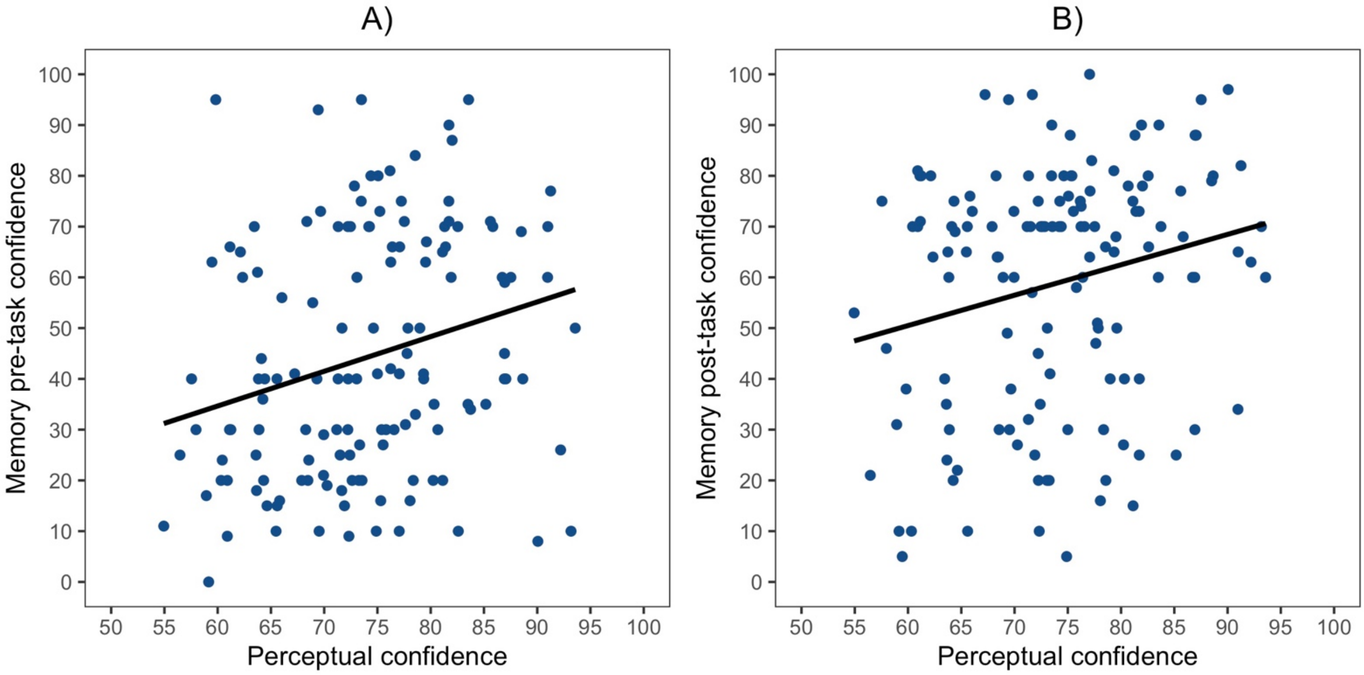


Note. Figures depicting correlations between, A) mean confidence in the number and contrast tasks, B) metacognitive efficiency in the number and contrast tasks, and C) post-task confidence ratings in the number and contrast tasks.

memory tasks, Pearson's correlations were conducted on the collapsed scores detailed above and those derived from the memory task. We found that mean perceptual confidence was significantly correlated with pre-task (r(136) = .25, p = .002) and post-task (r(136) = .22, p = .008) memory confidence where participants who displayed higher confidence in their perceptual performance also displayed higher confidence in their memory performance (see Figure 11A and Figure 11B, respectively). Although there was a trend towards a negative correlation between mean perceptual confidence and reminder bias (r(136) = -.14, p = .057, one-tailed as specified in the pre-registration), it did not pass the conventional threshold for statistical significance. There was also no significant

Figure 11

Relationship between Memory and Perceptual Confidence



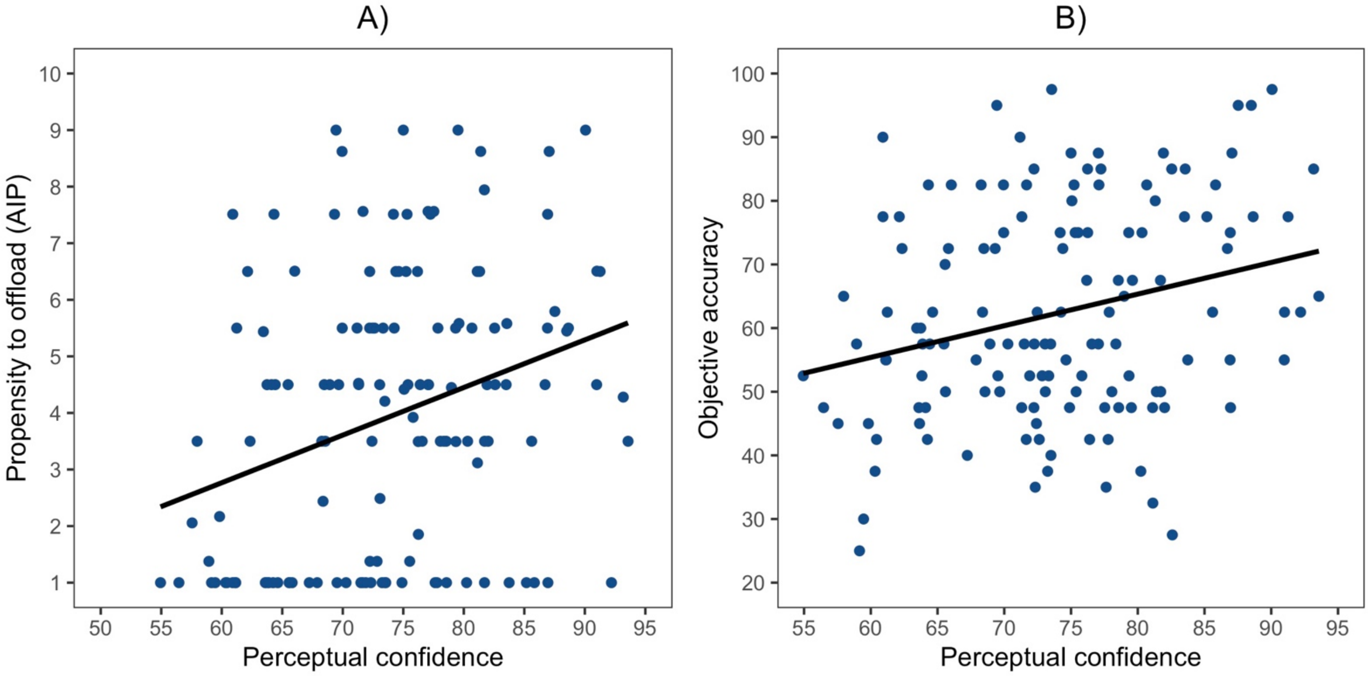
Note. Graphs illustrating, A) correlation between perceptual confidence and pre-task confidence, and B) correlation between perceptual confidence and post-task memory confidence.

correlation between post-task perceptual confidence and reminder bias ($r(136) = -.08$, $p = .17$).

Interestingly, both mean perceptual confidence ($r(136) = .30$, $p < .001$) (see Figure 12A) and post-task perceptual confidence ratings ($r(136) = .28$, $p < .001$) were positively correlated with AIP which is an inverse measure of how often participants choose to set reminders in the memory task. This shows that participants with higher confidence in their perceptual judgements not only tended to predict better performance in the memory task, but also set fewer reminders in the memory task. Furthermore, mean perceptual confidence ($r(136) = .26$, $p = .002$) (see Figure 12B) and post-task perceptual confidence ($r(136) = .35$, $p < .001$) were significantly correlated with forced internal accuracy in the intention offloading

Figure 12

Relationship between Perceptual Confidence and Measures in the Memory Task



Note. Graphs showing, A) the relationship between perceptual confidence and participants' propensity to offload (i.e., AIP), and B) the relationship between perceptual confidence and objective accuracy in the memory task.

task. So even though perceptual confidence was not correlated with perceptual accuracy (which was stabilized using a staircase procedure), participants with higher confidence in their perceptual confidence judgements had greater accuracy in the memory task.

As per the pre-registration the correlation coefficient derived from the association between mean perceptual confidence and reminder bias, and the correlation coefficient derived from the association between pre-task memory confidence and reminder bias were then compared to investigate whether one was significantly more predictive of reminder bias than the other. The difference between these two correlation coefficients was not significant (z = -.69, p = .49).

Seeing as a significant correlation between perceptual confidence and AIP was found, an additional analysis that was not pre-registered was conducted. This analysis compared the correlation coefficient derived from the correlation between mean perceptual confidence and AIP and the correlation coefficient derived from the correlation between pre-task memory confidence and AIP. A significant difference between these two correlation coefficients was not found (z = -.10, p = .92). Therefore, there was no evidence for any contribution of task-specific metacognitive signals over and above domain-general ones.

An additional analysis comparing the correlation between perceptual confidence and AIP, and perceptual confidence and reminder bias was also conducted. This was significant (z = -2.88, p = .004) adding further support for a greater influence of domain-general metacognitive signals on AIP (i.e., propensity to set reminders) than reminder bias (i.e., preference for reminders, relative to the optimal strategy).

As predicted, metacognitive efficiency in the perceptual tasks did not correlate with any of the measures derived from the memory task (rs < .13, ps > .14).

Domain-general versus domain-specific metacognitive signals

To investigate whether reminder bias was influenced by both domain-general and domain-specific confidence signals, a multiple linear regression with reminder bias as the dependent variable and perceptual confidence, and pre-task memory confidence as independent variables was conducted. The perceptual confidence measure was generated by transforming the mean trial-by-trial confidence and post-task confidence ratings into Z scores. These scores were then collapsed across the two measures. There was a significant effect of pre-task memory confidence (β = -.02, SE = .008, t(135) = -2.18, p = .03) but not perceptual confidence (β = -.18, SE = .22, t(135) = -.80, p = .43). Since a Pearson's correlation also did not show a significant correlation between perceptual confidence and reminder bias, bias towards/away from reminders is likely dominated by domain-specific metacognitive signals.

Since there was a significant positive correlation between perceptual confidence and AIP, an additional multiple linear regression that was not included in the pre-registration was conducted. In this model, AIP was included as the dependent variable and pre-task memory confidence rating and perceptual confidence were included as independent variables. This analysis evaluated whether the use of reminders was related to domain-specific along with the domain-general confidence found above. In this model, both pre-task memory confidence rating (β = .03, *SE* = .008, *t*(135) = 3.02, *p* = .003) and perceptual confidence (β = .66, *SE* = .23, *t*(135) = 2.94, *p* = .004) were significant, suggesting that AIP was influenced by both domain-general and domain-specific confidence signals.

Discussion

The present chapter examined the link between reminder-setting behaviour in a memory task and domain-general versus task-specific confidence signals influencing this behaviour. With regards to the memory task, the findings of previous research were replicated where it was found that decisions of whether or not to set reminders are linked to participants' metacognitive evaluations (see Gilbert et al., 2020; Kirk et al., 2021). Therefore, confidence signals play an important role in cognitive offloading.

With regards to the perceptual tasks, it was found that metacognitive efficiency and confidence in one perceptual task was related to its analogous measure in the second perceptual task. These findings support those of Gilbert (2015b Experiment 2) and Song et al. (2011) who found that, 1) metacognitive sensitivity correlates with its corresponding measure in perceptual tasks, 2) confidence correlates with its corresponding measure in the perceptual tasks, and 3) metacognitive sensitivity and confidence are not correlated with each other in perceptual tasks.

Domain-general account of metacognitive confidence

With regards to cross-domain associations, evidence for a domain-general component of metacognitive confidence was found where confidence in perceptual tasks was associated with memory confidence even though perceptual confidence was decorrelated from perceptual accuracy using a staircase procedure (see also Gilbert, 2015b; Mazancieux et al., 2020; McCurdy et al., 2013). This supports the domain-general account of metacognition which proposes that individuals use domain-general metacognitive signals when evaluating their performance (de Gardelle & Mamassian, 2014; Faivre et al., 2017). Furthermore, there was also a cross-domain link between perceptual confidence and reminder-setting in the memory task (i.e., AIP) where participants with lower perceptual confidence tended to set more reminders in the memory task.

The main novelty of this study was that in addition to participants' overall level of reminder-setting behaviour, their bias towards/away from setting reminders relative to optimal strategy was also investigated. As in earlier studies (see Gilbert et al., 2020; Kirk et al., 2021), both these measures were associated with memory confidence. However, unlike the overall index of reminder-setting behaviour (i.e., AIP, *propensity* to offload), there was no significant correlation between perceptual confidence and reminder bias (i.e., *preference* to offload). Furthermore, the correlation between perceptual confidence and AIP was significantly greater than the correlation between perceptual confidence and reminder bias, substantiating the cross-domain confidence effects on AIP, but not reminder bias.

This finding could be explained by some of the results in the current chapter. Before doing this, it should be pointed out that the correlation between perceptual confidence and reminder bias was marginally significant and that it is difficult to interpret null results. However, one possible explanation for this result could be that even though perceptual confidence was dissociated from perceptual accuracy, it was still correlated with objective accuracy in the memory task. This suggests that perceptual confidence cannot be considered a “pure” measure of metacognitive bias as it relates to cognitive ability in some way. Thinking of metacognition as an inferential process rather than simply a read-out of cognitive performance (see Koriat, 2007 for a review), this could be considered a rational strategy. Seeing as performance does tend to correlate across tasks (the *g* factor), on average, it is rational for an individual whose performance is high in one domain to predict higher performance in another domain. So, an individual with relatively good memory performance might also predict relatively good perceptual performance, even though the latter was stabilized by the staircase procedure. This could explain the link between perceptual confidence and memory accuracy.

One consequence of this link is that it would eliminate the correlation between perceptual confidence and reminder bias as people with low confidence in their perceptual judgements tend to set more reminders in a memory task simply because they might also *need* more reminders (seeing as low perceptual confidence is linked to low memory performance). Therefore, the cross-domain link is clearer for the *propensity* to offload than it is for reminder bias.

Another possible explanation for this comes from the results of Ball et al. (2021). They found a similar result where in their study they used three different version of the same memory task and found that even though the paradigm was the same, reminder bias only correlated with memory confidence in its own task but not with memory confidence from a different task. This result along with those presented in this chapter indicate that reminder bias might be influenced by task-specific confidence signals.

Metacognitive efficiency

The results presented in this chapter support the domain-specific account of metacognitive efficiency where it correlated with its analogous measure in the perceptual tasks but did not correlate with any of the measures in the memory task. This finding supports that of previous research where the majority of results have concluded in favour of a domain-specific account for metacognitive efficiency (e.g., Baird et al., 2013, 2015; Gilbert, 2015b Experiment 2; Morales et al., 2018). This is also corroborated by neuropsychological results where perceptual metacognitive sensitivity has been found to be related to the anterior prefrontal cortex (aPFC) (Baird et al., 2013, 2015; McCurdy et al., 2013) and lesions to the aPFC have shown to selectively impair perceptual sensitivity while leaving memory task performance intact.

In this study, although metacognitive efficiency and confidence were correlated with their analogous measures in the two perceptual tasks, metacognitive efficiency in the Number task was also associated with confidence in the Number task and the Contrast task. This finding is surprising as these two measures are thought to be independent of each other (see Galvin et al., 2003; Song et al., 2011). However, the results of this chapter support those of Shekhar and Rahnev (2020) who found that metacognitive efficiency *depended* on confidence level, where metacognitive efficiency becomes less reliable for higher confidence criteria. So, in some samples, it is possible to find a correlation between metacognitive efficiency and metacognitive confidence (Shekhar & Rahnev, 2020).

Conclusion

In conclusion, this chapter found evidence for the influence of both domain-general and task-specific confidence signals on reminder-setting behaviour. However, these domain-general signals did not fully explain metacognitive influence on reminder-setting behaviour especially with respect to reminder bias which did not correlate with confidence across domains. There was also no evidence for a link between perceptual metacognitive efficiency and reminder-setting behaviour. These results suggest that although metacognitive interventions could have some cross-domain influence on offloading behaviour, task-specific interventions are more likely to have a stronger impact especially when it comes to preference to offload.

CHAPTER 5. CONSEQUENCES OF COGNITIVE OFFLOADING

Chapters two to four have explored factors (specifically metacognition and effort-minimization) influencing intention offloading. Chapter three found that effort-minimization influences cognitive offloading where incentivizing individuals with performance-based rewards reduces their bias towards using reminders. One explanation for this could be that remembering intentions internally is effortful and involves cognitive processes that are limited in capacity and are potentially applicable to a wide range of tasks across domains (Kurzban et al., 2013). Taking this argument into account, perhaps individuals offload information to conserve cognitive processes so that these processes can be redirected towards subsequent tasks, improving their performance.

A different explanation (termed the "law of less work" Hull, 1943) is that individuals have an intrinsic drive to avoid using internal memory resources and to instead rely on external resources (Ballard et al., 1997a). Related to this point, arguments have been raised that relying on external memory can impair our internal cognitive resources (Baldwin et al., 2011). This negative effect of using external resources was investigated by Sparrow et al. (2011) who found that offloading information can make it more difficult to remember that information. However, offloading information might free cognitive resources that could be redirected towards other tasks. Therefore, the current chapter wanted to investigate the consequences of offloading one set of material on subsequently presented material.

Saving-enhanced memory effect

Previous research investigating the consequences of saving the contents of one computer file on memory for the contents of another file has found a memory enhancement for the contents of both the file that is saved and the contents of another file

(Runge et al., 2019; Storm & Stone, 2015). For example, Storm and Stone (2015) conducted an experiment where participants were first presented with two lists of eight words, and then their memory for both lists was tested. On some trials, participants were allowed to "save" the first list (Save trial) while on other trials they were forced to remember both lists using unaided memory (No Save trial). Storm and Stone (2015) found that, 1) saving and restudying a file led participants to remember a higher proportion of information from that file than when it was not saved, and 2) saving a file before studying a new file significantly improved recall of the contents of the new file. The latter finding suggests that offloading previous material facilitated the encoding and remembering of new information.

This result was replicated by Runge et al. (2019) who found that the benefits of offloading memory onto external resources was not just limited to memory performance but also improved performance on a subsequent task by redirecting cognitive resources. In another replication, Runge et al. (2020) demonstrated that the saving-enhanced memory effect held (and was even stronger) when using motor sequences instead of using word lists.

Together, these findings suggest that there might be two possible mechanisms that contribute to the saving-enhanced memory effect. The first is enhanced encoding where saving information reduces memory load allowing cognitive resources to be redirected towards subsequent tasks. The second is reduced interference at recall where saved information can be temporarily forgotten (as it can be accessed at a later time point), thus reducing interference on subsequently encoded information.

Both these processes align with the results typically found in list method directed forgetting (LMDF) literature where cueing participants to forget a previously studied list (List A) and to remember a new list instead (List B) leads to the forgetting of the first list and a memory enhancement for the second list (see Bäuml et al., 2010; Sahakyan et al., 2013 for a review). However, within the domain of cognitive offloading, the studies discussed above (Runge et al., 2019; Storm & Stone, 2015) always involved either saving or not saving List A. Therefore, the first aim of the current chapter was to adapt the paradigm used by Storm and Stone (2015) to manipulate which of the two lists (List A or List B) is offloaded in order to investigate whether the saving-enhanced memory effect is sustained even when the to-be-remembered information is presented before the saved information (i.e., List B is saved instead of List A).

Directed forgetting

Research in LMDF has provided contradicting results for when participants have been cued to forget the second list. Sahakyan (2004) found that attempting to forget List B had both direct and indirect costs (also see Racsmány et al., 2019). In their experiment, Sahakyan (2004) presented participants with three lists of words which they had to study and subsequently recall. They found that forgetting the middle list (List B in this case) led to reduced recall not only for List B but also for List A, even though List A was not intended for forgetting. This effect was found even when the lists consisted of separate distinct categories. In contrast, Kliegl et al. (2013) found that participants were able to selectively forget List B items without forgetting List A items regardless of the modality of item presentation (visual presentation versus auditory presentation and the discriminability between the two lists (relevant versus irrelevant) information. This result was replicated in both short and long lists (Kliegl et al., 2020b).

The contrasting results between the two experiments could be explained by one difference between the studies. This difference relates to the order in which the tests were presented to the participants. While in Sahakyan's (2004) experiment, participants had to recall List A before recalling List B, in Kliegl et al.'s (2013) experiment, this order was counterbalanced between participants. This notion was explored by Pastötter et al. (2012) who found that reliable List B memory enhancement arose only when List B was recalled first. This suggests that testing List A first might reinstate proactive interference by re-exposure to List A material, subsequently leading to a reduction in List B enhancement. List A forgetting, however, was found regardless of which list was recalled first even though participants recalled more List A items when these were tested first (see also Aguirre et al., 2020 for a similar result). Therefore, the second aim of this chapter was to manipulate the order in which participants recalled the contents of the two lists. That is, half of the participants were asked to recall List A first and List B second while for the other half this test order was reversed.

Experiment 1

The aims of the first experiment in this chapter were twofold. The first was to compare the benefits of offloading List A versus offloading List B. This was done by adapting the paradigm used by Storm and Stone (2015) where participants were asked to study two lists of eight words on which they were then tested. Participants were asked to either save List A, save List B or to perform the task using unaided memory.

The second aim was to investigate whether the saving-enhanced memory effect on List B recall when List A is saved is affected by the order in which the two lists are tested. To examine this, half of the participants were asked to recall List A items before List B items and for the other half of the participants, this test order was reversed where they were asked to recall List B items before List A items.

This experiment evaluated the following hypotheses:

- If saving a file works as a forget cue, an increase in the recall for List B items on *List A-Saving* trials will be found and reduced recall enhancement for List A items on *List B-Saving* trials will be found.
- On *List A-Saving* trials, the saving-enhanced memory effect for List B will be larger when List B is recalled first than when List A is recalled first, as found for forget cues in LMDF literature.
- On *List A-Saving* trials, the proportion of words recalled from List A will be larger when it is recalled second than when it is recalled first. This is because List B would have already been tested and participants would be able to focus completely on List A when restudying it.

Before commencing data collection, we preregistered our hypotheses, experimental procedure and analysis plan (https://osf.io/jxyqs/).

Method

Participants

A total of 102 participants (51 in each group) (*mean age* = 35.51 years; *SD age* = 11.82 years; *range* = 18-27; 45 male; 56 female; 1 other) were recruited through Prolific (https://www.prolific.co). Participation was restricted to volunteers aged at least 18 years who spoke English as their first language. Ethical approval for this study was granted by UCL Research Ethics Committee (1584/003) and participants provided informed consent before participating in the study. Participation took approximately 40 minutes and participants were paid £5 as compensation.

To estimate sample size, a statistical power analysis was conducted using G*Power 3.1 (Faul et al., 2007). The power calculation for this experiment was based on the results of Storm and Stone (2015 Expeirment 3). In their experiment, two lists of eight words were presented to participants and they found a saving-enhanced memory effect for recall of List B when List A was saved. The effect size (d_z) for this analysis was .93. To find an effect on List A recall when list B is saved, a more conservative approach was used where the previous number (d_z = .93) was halved. This resulted in an effect size (d_z) of .465. To achieve 90% power to replicate an effect of this size (two-tailed test, α = .05), a sample of 51 participants was required. Since the test order factor in this experiment was between-subjects, the total sample for this experiment was 102 participants with 51 participants in each test order group.

Participants whose memory performance (averaged across conditions) exceeded 3 median absolute deviation units (MAD; Christophe Leys et al., 2013) (outliers) were excluded (n = 8). Furthermore, 3 participants reported cheating (e.g., writing things down or taking pictures) and were also excluded. All participants excluded in this experiment (n = 11) were replaced so that the final sample totalled 102 participants (51 in each group).

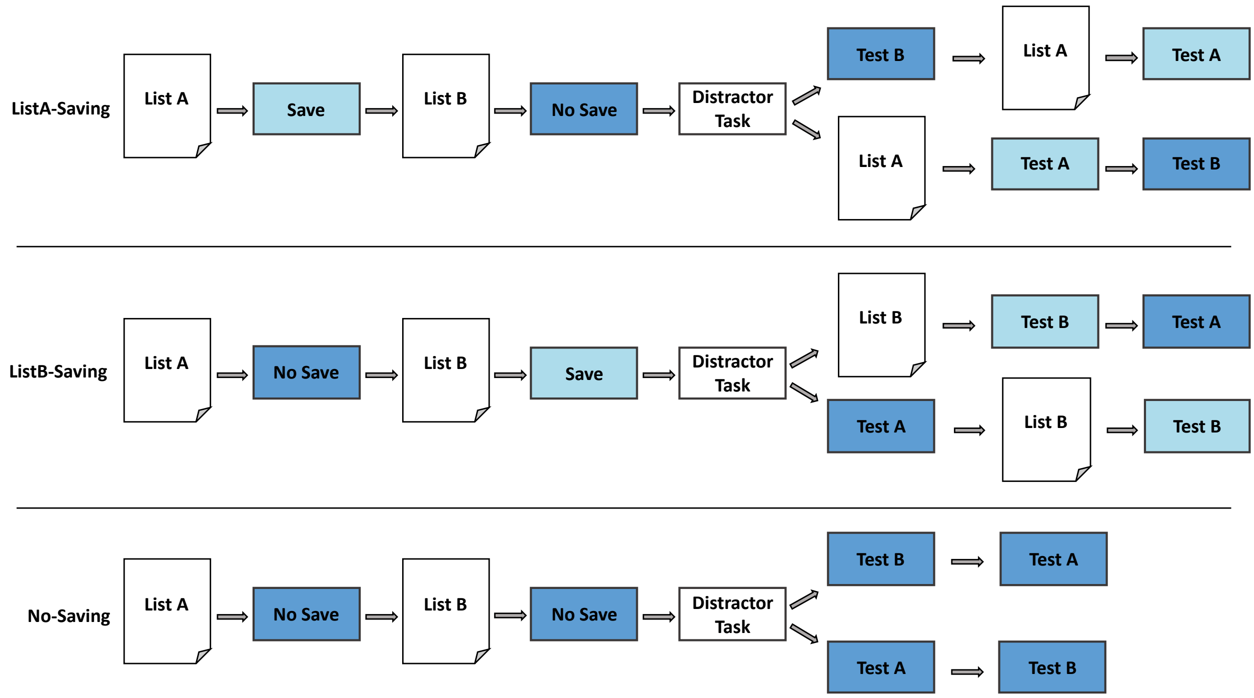
Design

This task was programmed using Gorilla (https://gorilla.sc/) and had two manipulations. The first was whether and which list was saved and restudied. This was manipulated within-subjects where each participant performed four trials saving List A (List A-Saving trials), four trials saving List B (List B-Saving trials), and four trials using unaided memory where neither list could be saved (No-Saving trials).

The second manipulation concerned the order in which the two lists were tested at the end of each trial. This was manipulated between-subjects where half of the participants were tested on List B first (as in Storm & Stone, 2015) and the other half were tested on List A first (see Figure 13 for a schematic representation of the task).

Figure 13

Schematic Representation of the Task



Note. Participants were instructed to study both List A and List B. On List A-Saving trials, participants saved List A before studying List B; on List B-Saving trials, participants saved list B before the distractor task; on No-Save trials, participants did not save any list. After a short distractor task, participants were tested on the two list. Half of the participants were tested on List A first while the other half were tested on List B first. On List A-Saving and List B-Saving trials, participants could restudy the saved list before being tested on it.

Materials

For this experiment, 192 common nouns (4 to 7 letters long) were selected from the Paivio Word List Generator (http://euclid.psych.yorku.ca/shiny/Paivio/). For each participant, 24-word lists were randomly generated (two lists consisting of List A and List B for each trial). Both lists consisted of eight words each. After each trial participants played a “Spot the difference” puzzle where they had to find 10 differences between two similar images. These pictures were taken from “La Settimana Enigmistica” – Italy.

Procedure

First, participants provided informed consent and were randomly assigned to one of the test order conditions. They were then informed that on each trial of the task they would have to study the contents of two lists and that they would be tested on the contents of those lists at the end of each trial. Participants completed three practice trials where they were introduced to each of the three conditions (No-Save, List A-Saving, and List B-Saving) individually. After completing the practice trials for each condition, participants were introduced to the break task (see below) and were presented with a short comprehension quiz which tested their knowledge on the different elements of the task. For every mistake that they made, they received further clarification on what that component of the task entailed. This was done to ensure that participants were confident in their knowledge of what to do in the task before beginning the main experimental trials.

Participants then completed 12 trials of the main task. In the main task, a third of the trials were List A-Saving trials, a third of the trials were List B-Saving trials and the remaining third of trials were No-Save trials. These trials were presented in randomized order. On each trial, participants studied List A and List B for 15 seconds each. On List A-Saving trials, participants were instructed to save List A after studying it. To save this list, they were prompted to press a button labelled “Save” on the screen. On List B-Saving trials, participants were instructed to save List B after studying it. To save this list, like List A-Saving trials, they were prompted to press a button labelled “Save” on the screen. On the No-Save trials, participants were not allowed to save or restudy either list. Participants were informed that saving a list would ensure that they would be able to restudy it prior to test.

After the study phase, there was a short 20 second delay during which participants were asked to count backwards by threes from a three-digit number between 200 and 999 (as in Storm & Stone, 2015). When the time ran out, they were prompted to type in the last number that they had reached in the sequence.

After completing the digit task, participants were presented with two recall tests where they were asked to recall the words presented on List A and List B. Half of the participants were asked to recall List A items first and List B items second while for the other half this testing order was reversed. For the recall test, participants had 45 seconds to type all the words they could recall from the instructed list. On List A-Saving trials, participants were instructed to restudy List A for 15 seconds before they were tested on it. Similarly, on List B-Saving trials, participants were instructed to restudy List B for 15 seconds before they were tested on it. After each trial, participants were given a “Spot-the-difference” puzzle as a distractor task for one minute before beginning the next trial.

At the end of the experiment, participants were thanked for their time, paid and debriefed (a demonstration of the experiment can be found at: https://app.gorilla.sc/openmaterials/220207).

Results

Recall performance for List A

Analyses were conducted using R (version 4.0.3) and analyses were conducted as per the pre-registration plan.

To investigate the proportion of words correctly recalled from List A, a 3 x 2 mixed ANOVA with within-subjects factor Saving Condition (List A-Saving vs. List B-Saving vs. No-Saving) and between-subjects factor Test Order (Test A first versus Test B first) was conducted. Although the main effect of Test Order was not significant ($F(1,100) = 2.35$, $p = .13$, $\eta^2_p = .023$), the main effect of Saving Condition ($F(1,100) = 280.46$, $p < .001$, $\eta^2_p = .737$) and the interaction $F(1,100) = 11.36$, $p < .001$, $\eta^2_p = .102$) were significant.

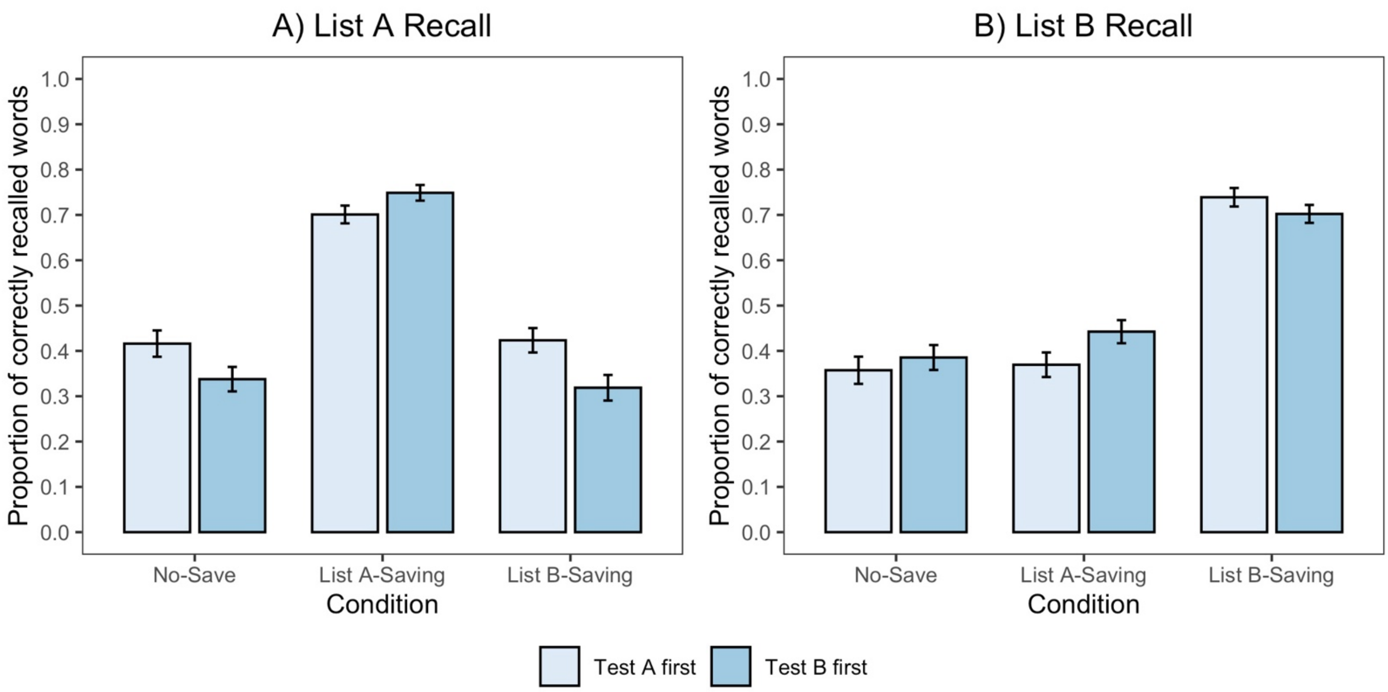
These results were then further qualified with a series of *t*-tests. First, three independent samples *t*-tests were conducted. It was found that on List A-Saving trials, List A recall did not differ significantly when it was tested second (M = .75, SD = .12) compared to when it was tested first (M = .70, SD = .21) (t(100) = 1.83, p = .07, d = .36). In contrast, List A recall was significantly higher when it was tested first in the List B-Saving condition (M = .42, SD = .19) compared to when it was tested second (M = .32, SD = .20) (t(100) = 2.69, p = .008, d = .53), and marginally significant when it was tested first in the No-Saving condition (M = .42, SD = .21) compared to when it was tested second (M = .34, SD = .19) (t(100) = 1.98, p = .051, d = .39) (see Figure 14A).

Three paired *t*-tests were then computed separately for the two test order groups. In the group where List A was tested first, there was a significant difference between the No-Save condition and the List A-Saving condition (t(50) = 10.55, p < .001, d_z = 1.48), and between the List A-Saving condition and the List B-Saving condition (t(50) = 12.34, p < .001, d_z = 1.73) where recall of List A was higher in the List A-Saving condition than in the No-Save or List B-Saving condition. There was no significant difference in the proportion of List A words recalled in the No-Save and the List B-Saving condition (t(50) = 0.41, p = .685, d_z = 0.06) suggesting that saving List B did not improve recall of List A words even when List A was recalled first.

In the group where List B was tested first, there was a significant difference between the No-Save and List A-Saving conditions (t(50) = 155, p < .001, d_z = 2.1) and the List A-Saving and List B-Saving conditions (t(50) = 13.81, p < .001, d_z = 1.93). Recall of List A was higher in the List A-Saving condition than in the List B-Saving and No-Save conditions. Once again, no significant difference between the List B-Saving and No-Save conditions was found (t(50) = 1.25, p = .218, d_z = 0.18).

Figure 14

Recall of List A and List B as a Function of Saving Condition and Test Order



Note. Error bars represent standard errors of the mean.

Recall performance for List B

To investigate the proportion of words recalled from List B, a 3 x 2 mixed ANOVA analogous to the one above with within-subjects factor Saving Condition (List A-Saving vs. List B-Saving vs. No-Saving) and between-subjects factor Test Order (Test A first versus Test B first) was conducted. Similar to the results for List A, no significant main effect of Test Order was found ($F(1,100) = .54$, $p = .47$, $\eta^2_p = .005$) but a significant main effect of Saving Condition ($F(1,100) = 230.35$, $p < .001$, $\eta^2_p = .697$) and a significant interaction ($F(1,100) = 4.73$, $p = .014$, $\eta^2_p = .045$) were found.

These results were further qualified with a series of *t*-tests. First, three independent samples *t*-tests were conducted. It was found that on List A-Saving trials, recall of List B was slightly higher when it was tested first (M = .44, SD = .18) compared to when it was tested second (M = .37, SD = .20) but this result did not reach the conventional threshold for significance ($t(100) = 1.96$, $p = .053$, $d = .39$). On List B-Saving trials, there was no difference between when List B was tested first (M = .74, SD = .15) compared to when it was tested second (M = .70, SD = .14) ($t(100) = 1.29$, $p = .199$, $d = .26$). Similarly, on No-Save trials, there was no difference between when List B was tested first (M = .36, SD = .21) compared to when it was tested second (M = .39, SD = .20) ($t(100) = 0.69$, $p = .500$, $d = .14$) (see Figure 14B).

Next, three paired t-tests were computed separately for the two test order groups. In the group where List A was tested first, there was a significant difference between the List A-Saving and List B-Saving conditions ($t(50) = 14.26$, $p < .001$, $d_z = 2.00$) and the No-Save and List B-Saving conditions ($t(50) = 11.53$, $p < .001$, $d_z = 1.61$) where recall of List B was higher in the List B-Saving condition than in the List A-Saving or No-Save conditions. However, there was no difference between the List A-Saving and No-Save conditions ($t(50) = 0.53$, $p = .60$, $d_z = .07$).

In the group where List B was tested first, there was a significant difference between the List A-Saving and List B-Saving conditions ($t(50) = 10.28$, $p < .001$, $d_z = 1.44$) where List B recall was higher in the List B-Saving condition than in the List A-Saving condition. Furthermore, there was a significant difference between the List B-Saving condition and the No-Save condition ($t(50) = 12.69$, $p < .001$, $d_z = 1.78$) where List B recall was higher in the List B-Saving condition than in the No-Save condition. Furthermore, a significant difference between the List A-Saving condition and the No-Save condition ($t(50) = 3.27$, $p = .002$, $d_z = .46$) was also found where List B recall was higher when List A was saved than when neither list was saved. These results suggest that saving List A only enhanced recall of list B when List B was tested first.

Recall performance for the list that was saved

To investigate recall performance for the list that was saved (i.e., recall performance of List A when List A was saved and recall performance of List B when List B was saved), a 2 x 2 mixed ANOVA with within-subjects factor Saving Condition (List A-Saving vs. List B-Saving) and between-subjects factor Test Order (Test A first versus Test B first) was conducted on the proportion of correctly recalled words from the list that was saved and restudied. There was no significant main effect of Saving Condition ($F(1,100) = 0.14$, $p = .70$, $\eta_p^2 = .001$) and Test Order ($F(1,100) = 0.05$, $p = .82$, $\eta_p^2 < .001$) but there was a significant interaction ($F(1,100) = 13.62$, $p < .001$, $\eta_p^2 = .12$).

To qualify the interaction, paired *t*-tests were conducted separately for the two test order groups. In the group where List A was tested first, recall was significantly better for List B when List B was saved ($M = .74$, $SD = .15$) than it was for List A when List A was saved ($M = .70$, $SD = .14$) ($t(50) = -2.16$, $p = .04$, $d_z = .30$).

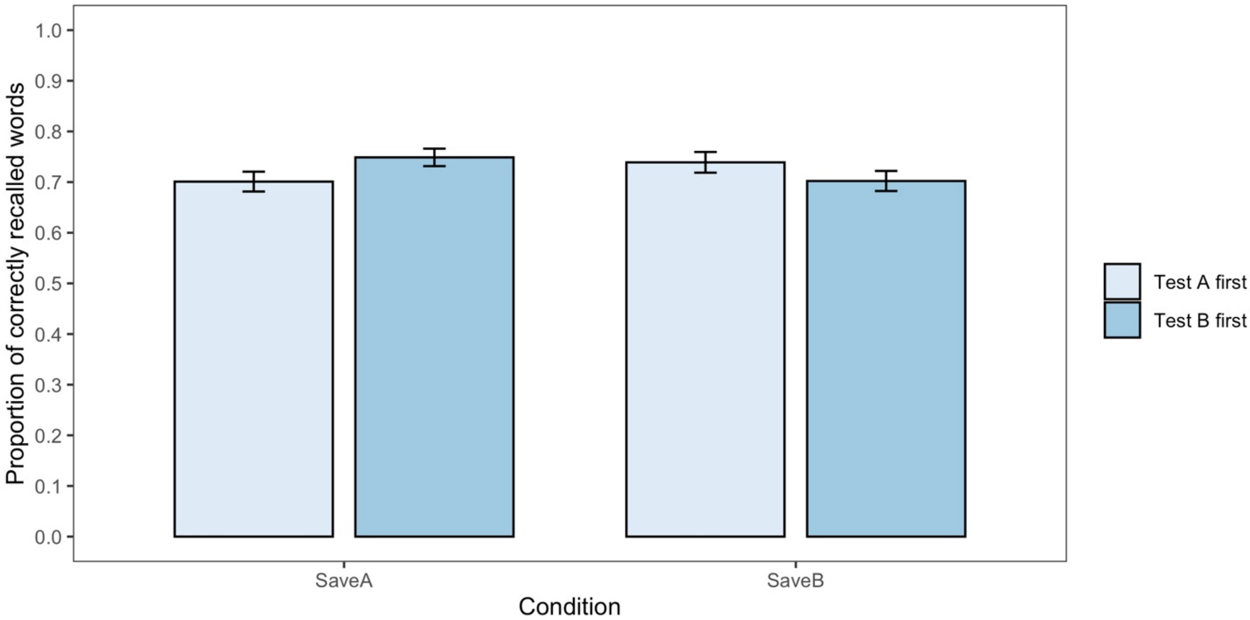
Conversely for the group where List B was tested first, recall was significantly better for List A when List A was saved ($M = .75$, $SD = .12$) than it was for List B when List B was saved ($M = .70$, $SD = .14$) ($t(50) = 3.16$, $p = .003$, $d_z = .44$). Therefore, in both groups, performance was better for whichever list was restudied and tested second (see Figure 15).

Recall performance for the list that was not saved

To investigate recall performance for the list that was not saved (i.e., recall performance of List A when List B was saved and recall performance of List B when List A was saved), a 2 x 2 mixed ANOVA with within-subjects factor Saving Condition (List A-Saving vs. List B-Saving) and between-subjects factor Test Order (Test A first versus Test B first) was

Figure 15

Recall of List A and List B when Saved



Note. Error bars represent standard errors of the mean

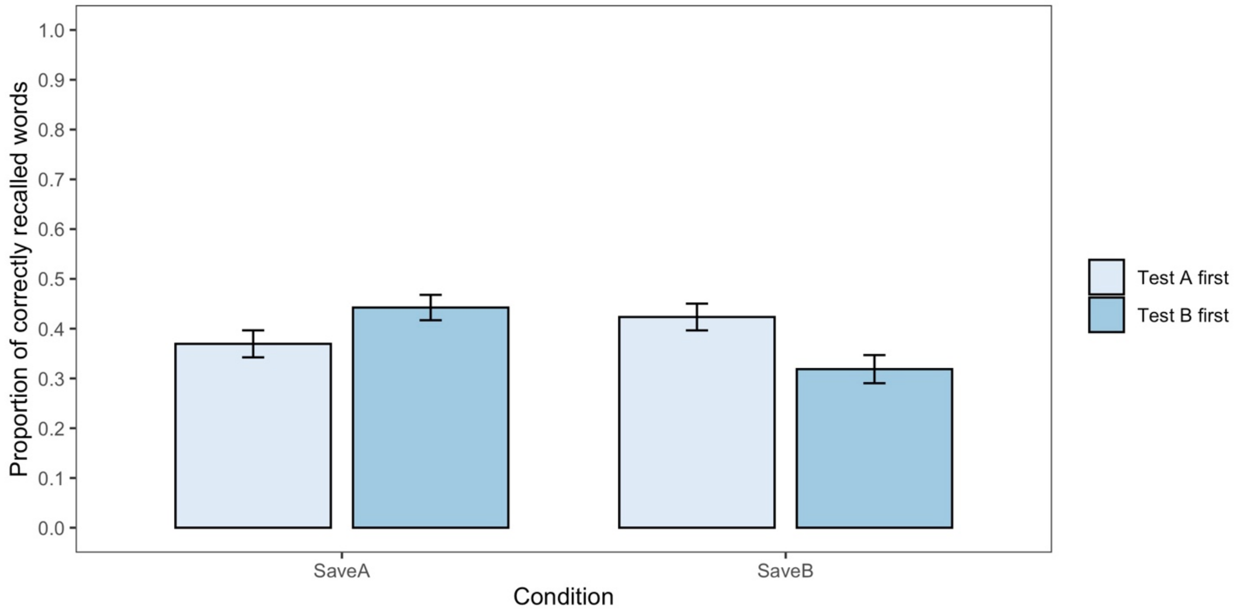
conducted on the proportion of correctly recalled words from the list that was not saved. There was no significant main effect of Test Order (F(1,100) = 0.22, p = .64, η_p^2 = .002). But there was a significant main effect of Trial Condition (F(1,100) = 4.45, p = .04, η_p^2 = .043) where recall was better for List B when List A was saved (M = .41, SD = .19) than for List A when List B was saved (M = .37, SD = .20). The interaction was also significant (F(1,100) = 28.81, p < .001, η_p^2 = .22).

To qualify the interaction, paired t-tests were conducted separately for the two test order groups. In the group where List A was tested first, the proportion of words recalled from List A when List B was saved was higher (M = .42, SD = .19) than the proportion of words recalled from List B when List A was saved (M = .37, SD = .19) (t(50) = 2.49, p =.02, d_z = .35).

Conversely, in the group where List B was tested first, the proportion of words

Figure 16

Recall of List A and List B when Not Saved



Note. Error bars represent standard errors of the mean.

recalled from List B when List A was saved was higher (*M* = .44, *SD* = .18) than the proportion of words recalled from List A when List B was saved (*M* = .32, *SD* = .20) ($t(50)$ = 4.94, $p < .001$, d_z = .69) (see Figure 16).

Discussion

Overall, the results of the present experiment found that saving and restudying List A or List B improved recall for the offloaded material. Furthermore, in line with the first hypothesis, a saving-enhanced memory effect was found for List B when List A was saved, but no benefit for List A items when List B was saved was found. This first set of results supports those of Storm and Stone (2015), who also found a saving-enhanced memory effect for List B items on List A-Saving trials. These results also support those of Sahakyan (2004) where there was no benefit to List A recall when participants were cued to forget List B. This held true regardless of test order (see also Pastötter et al., 2012 for similar results).

Furthermore, it was found that the saving-enhanced memory effect for List B was only found in the group of participants where List B was tested first. This result also supports the second hypothesis where it was predicted that on List A-Saving trials, the saving-memory enhancement for List B would be larger when List B was recalled first than when it was recalled second.

No support was found for the third hypothesis where it was predicted that there would be an effect of testing order on List A-Saving trials. However, it was found that in the group where List A was tested first, accuracy was higher for List B when it was saved than for List A when it was saved. The opposite was true for the group where List B was tested first.

Experiment 2

The results of Experiment 1 suggest that when presented with two pieces of material to study and when given the opportunity to offload one of the two, it is advantageous to offload the first piece of information and to rely on internal memory for the second piece of information. As a follow up, the second experiment in this chapter aimed to extend the findings of the first experiment to investigate whether when given free choice, participants would have a preference towards saving a specific list. Specifically, the present study examined whether participants would choose to save and restudy the most effective list, i.e., List A. This was done by adapting the paradigm used in Experiment 1. On each trial, participants were given free choice to either save List A, List B or to remember both lists using unaided memory. In line with Experiment 1, this experiment also investigated whether participants' preferences were influenced by the order in which the two tests were recalled. It was predicted that, 1) participants would prefer to save a list over using their unaided memory only, and that their preference for which list to offload would depend on test order, that is, 2) they would prefer to save List A when List B is tested first, and they would prefer to save List B when List A is tested first.

Previous literature has found that decisions to offload information onto external resources is influenced by one's confidence in their cognitive abilities where individuals are more likely to set a reminder when they have low confidence in their cognitive abilities (e.g., Boldt & Gilbert, 2019; Gilbert, 2015b; Gilbert et al., 2020). Therefore, the current study also aimed to investigate whether participants' preference would be related to their confidence in their ability to remember the list unaided. To do this, participants were asked to estimate their accuracy for List A and List B items in the different conditions before commencing the task and also at the end of the main task. It was further hypothesised that, 3) preference for saving a list over relying on unaided memory would be associated with participants' confidence in their ability to recall List A and List B using unaided memory, and 4) preference for saving a specific list would be associated with lower confidence for recalling that list when it was not saved than for recalling the other list when it was not saved. For example, the likelihood of saving List A would be positively associated with the difference between their confidence for List B recall when List A is saved and their confidence for List A recall when List B saved.

Method

Participants

A total of 88 participants (*mean age* = 39.58 years; *SD age* = 12.85 years; *range* = 20 – 71; 41 male; 44 female; 3 other) were recruited through Prolific (https://www.prolific.co). Participation was restricted to volunteers aged at least 18 years who spoke English as their first language. Ethical approval for this study was granted by UCL Research Ethics Committee (1584/003) and participants provided informed consent before participating in the study. Participation took approximately 60 minutes and participants were paid £7.50 as compensation.

To estimate sample size, a statistical power analysis was conducted using G*Power 3.1 (Faul et al., 2007). This study was powered to detect a medium effect size (d = 0.5) for the analysis of whether participants have a preference towards saving a specific list. To achieve 90% power to replicate an effect of this size (two-tailed test, α = .05), a total sample size of 88 with 44 participants in each test order group was required.

Four participants who reported cheating in the final questionnaire were excluded and replaced.

Design

This experiment was programmed using Gorilla (https://gorilla.sc/) and consisted of two manipulations. The first manipulation was whether and which list was saved and restudied where at the beginning of each trial participants had to make two consecutive choices. The first was whether they would like to save one list or save no list. If they chose to save a list, they got a second choice where they had to choose between saving List A or saving List B. On 50% of the save trials participants were able to restudy the list they had chosen to save before being tested on it. On the other 50% of the save trials, participants were not represented with the list and so were not able to restudy the list they had chosen to save before being tested on it. This was done to ensure that participants encoded both lists even if they decided to save one of the lists. This made the design of the current experiment comparable with the previous one.

The second manipulation concerned the order in which the lists were tested at the end of each trial. This was manipulated between-subjects where half of the participants were tested on List B before List A (as in Storm & Stone, 2015) while for the other half this order was reversed.

Materials

The same stimuli materials as Experiment 1 were used in this experiment.

Procedure

The procedure of this experiment was similar to Experiment 1. Where the procedure differed was at the beginning of each of the 12 trials where participants were given a choice to either save List A, save List B or to perform the task using unaided memory. Participants were informed that on save trials, they would be given the opportunity to restudy the saved list on only half of the trials.

Furthermore, participants' pre-task and post-task confidence ratings were collected where they had to estimate their accuracy in recalling, 1) List A when no list was saved, 2) List B when no list was saved, 3) List A when List A was saved, 4) List A when List B was saved, 5) List B when List B was saved, and 6) List B when List A was saved.

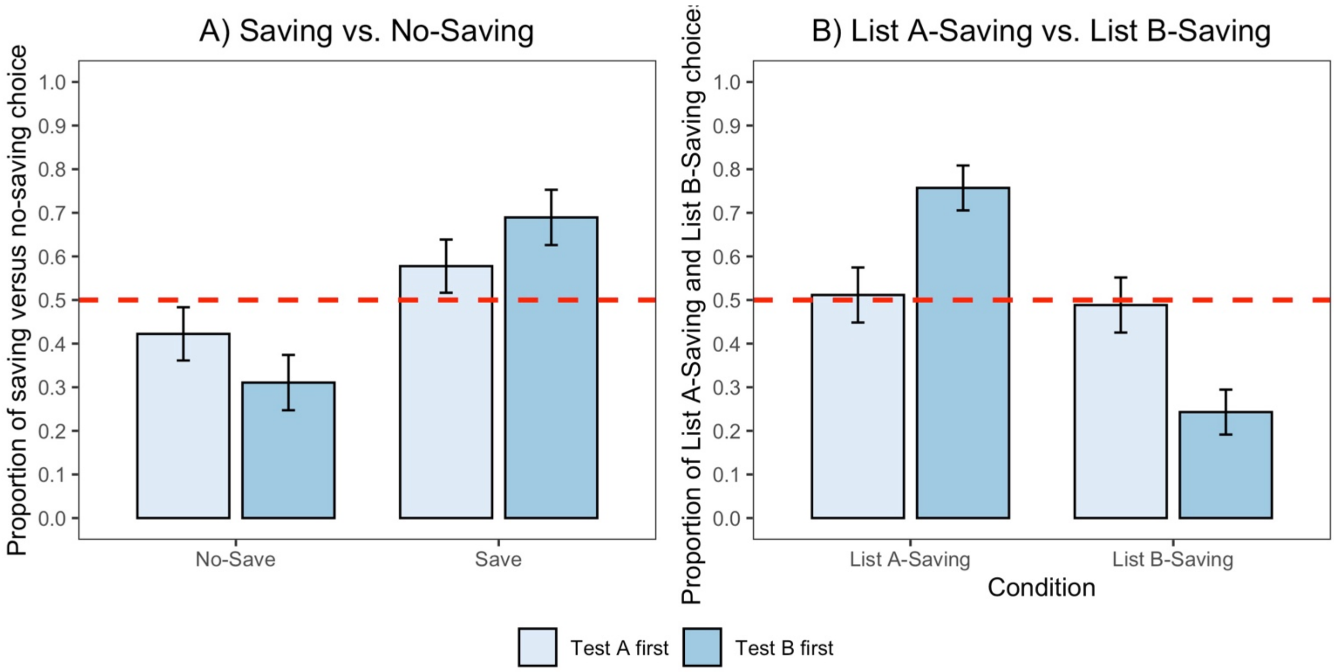
Results

Analyses were conducted as per the pre-registration plan. In addition to these, exploratory analyses on participants' post-task confidence ratings were also conducted.

First, whether participants had a preference towards saving a list over using their unaided memory was investigated. To investigate this, the proportion of times participants chose to save a list was computed and then this proportion was compared against ½ using a one-sample *t*-test seeing as participants could only choose one of two options (Save vs. No-Save). It was found that participants chose to save a list significantly more than half the time ($t(87) = 3.02$, $p = .003$, $d = .32$) showing a preference towards saving a list ($M = .63$, $SD = .41$) over using their unaided memory ($M = .37$, $SD = .41$) (see Figure 17A).

Figure 17

Participants' Offloading Choices



Note. Error bars represent standard error of the mean.

Participants' preference to offload a specific list was then examined. In other words, whether participants had a preference towards offloading either List A or List B was explored. For this analysis, all trials in which participants chose not to save a list were excluded. Furthermore, participants who never chose to save a list had to be excluded as they did not have any List A-Saving or List B-Saving trials. Then, the proportion of times participants chose to save List A over List B was computed and this proportion was compared against ½ using a one-sample *t*-test. This analysis found that participants chose to save List A significantly more than half the time ($t(74) = 3.08$, $p = .003$, $d = .36$) where participants showed a preference to save List A ($M = .63$, $SD = .37$) rather than saving List B ($M = .37$, $SD = .37$) (see Figure 17B).

Then, two independent samples *t*-tests were conducted. The first investigated if participants' preference towards saving a list versus using their unaided memory depended on the other in which the two tests were recalled. There was no significant difference between the proportion of times participants chose to offload a list when List A was tested first (M = .58, SD = .40) compared to when List B was tested first (M = .69, SD = .42) ($t(86)$ = 1.27, p = .21, d = .21). When controlling for participants' pre-task confidence ratings on their ability to remember List A and List B items using unaided memory in an ANCOVA model, these results did not change. More specifically, the main effect of test order ($F(1,84)$ = 1.34, p = .25, η^2_p = .02), pre-task confidence rating for ability to remember List A using unaided memory ($F(1,84)$ = .68, p = .41, η^2_p = .0008) and pre-task confidence rating for ability to remember List B using unaided memory ($F(1,84)$ = .02, p = .89, η^2_p = .0002) were all not significant.

The second independent samples *t*-test investigated if participants' preference to save List A versus List B depended on the order in which the two lists were recalled. The proportion of times participants chose to save List A was significantly larger when List B was tested first (M = .76, SD = .31) compared to when List A was tested first (M = .51, SD = .39) ($t(73)$ = 3.01, p = .004, d = .69). When controlling for the difference in participants' pre-task confidence ratings on accuracy for the two lists when the other list was saved (i.e., predicted recall for List B when List A was saved and predicted recall for List A when List B was saved), the results did not change. In this model, test order was significant ($F(1,72)$ = 8.86, p = .004, η^2_p = .11) but difference in pre-task confidence was not ($F(1,72)$ = .27, p = .61, η^2_p = .004) . When controlling for the difference in post-task confidence ratings instead, both test order ($F(1,72)$ = 5.05, p = .03, η^2_p = .066) and the difference in post-task confidence ($F(1,72)$ = 5.72, p = .02, η^2_p = .074) were significant.

Since the independent samples *t*-test investigating if participants' preference to save List A versus List B depended on the order in which the two lists were recalled was significant, two additional one-sample *t*-tests against ½ were conducted on the proportion of times participants chose to save List A in the two test order conditions. It was found that participants had a preference to towards offloading List A when List B was tested first ($t(36) = 4.99$, $p < .001$, $d = .82$) but not when List A was tested first ($t(37) = .18$, $p = .86$, $d = .03$).

In addition to the analyses above, the proportion of times participants saved List A (List B) was correlated with the difference between their pre-task confidence for List B recall when List A was saved and pre-task confidence for List A recall when List B was saved. This correlation was not significant ($r(73) = .07$, $p = .58$).

The same correlation was then repeated using post-task confidence measures. The proportion of times participants chose to save List A (List B) was significantly correlated with the difference between their post-task confidence for List B recall when List A was saved and post-task confidence for List A recall when List B was saved ($r(73) = .34$, $p = .003$) where List A (List B) saving increased as the difference between the two post-task confidence scores increased.

Discussion

Overall, the results of the present experiment found that when given a choice, participants preferred to save a list rather than use their unaided memory. This was in line with the first hypothesis. Furthermore, partial support was found for the second hypothesis where it was found that participants preferred to save the first list (List A) when List B was tested first but there was no preference to save List B when List A was tested first.

These results complement those of the first experiment where it was found that when given the opportunity to offload one of two lists, it is advantageous to offload the first list and rely on internal memory for the second list. Therefore, the results of this study suggest that when given free choice, participants choose to save and restudy the most effective list, i.e., List A (at least in this paradigm). However, this preference towards saving List A was only present when List B was tested first.

No support was found for the third hypothesis as neither the pre-task nor post-task confidence ratings of recalling List A or List B using unaided memory was correlated with preference for saving. Partial evidence for the fourth hypothesis was found where the difference between the post-task confidence ratings for List B recall when List A is saved and List A recall when List B is saved was associated with preference for saving List A/List B.

General discussion

The current chapter had four aims. Experiment 1 investigated, 1) if the saving-enhanced memory documented in previous research replicates when the to-be-remembered information is presented before the offloaded information (i.e., List B is saved instead of List A), and 2) if the saving-enhanced memory effect is affected by the order in which the materials are recalled. Experiment 2 evaluated, 1) whether, when given free choice, people have a preference for offloading rather than relying on internal memory and, 2) whether they choose to offload and restudy the most effective list (i.e., List A). A better understanding of factors that lead to a saving-enhanced memory effect and a better understanding of the factors that align individuals' offloading strategies with this effect is important to guide individuals towards the most beneficial offloading strategies.

In Experiment 1 it was found that the saving-enhanced memory effect was found for List B when list A was saved and tested second. But this effect was not sustained when List A was tested first. Experiment 1 also found that offloading List B did not enhance recall for List A. Experiment 2 found that when given a choice between offloading information and using unaided memory to remember information, participants chose to offload information. Moreover, it was also found that when choosing which list to save (List A or List B), participants preferred to save List A when it was tested second but did not have a preference towards saving one list over the other when List A was tested first.

Previous research has proposed two possible mechanisms associated with the saving-enhanced memory effect (Pastötter et al., 2012; Runge et al., 2020). The first mechanism is that saving information frees cognitive resources which would otherwise be utilised for rehearsing that information. Subsequently, these cognitive resources can then be utilised on other tasks thus improving performance. With regards to the current paradigm, saving either list could potentially free cognitive resources. However, in Experiment 1 it was found that saving List B did not lead to a saving-enhanced memory effect. This could be explained by the order in which the two lists were memorised. In both experiments, List A was always encoded before List B. So, when participants were presented with a List A-Saving trial, they knew that that list would be presented once again and so they could allocate their cognitive resources into encoding List B. On List B-Saving trials however, participants only know that List B can be saved after encoding both, List A and List B. So, on these trials, participants perhaps allocate resources to both lists which is why a saving-enhanced memory effect is not observed when List B is saved. This also supports the notion by Kurzban et al. (2013) who posited that remembering intentions internally is effortful and recruits cognitive processes that are limited in capacity.

The second mechanism is that saving information reduces proactive interference at recall as this information can be temporarily forgotten and accessed at a later time point, reducing interference for recall of subsequently encoded information (Pastötter et al., 2012). The findings of the first experiment showed that the saving-enhanced memory effect on List B was not sustained when List A was tested first. This finding supports the account of saved information reducing proactive information as testing List A might have reinstated proactive interference by re-exposing participants to the List A material, thus causing a reduction in subsequent List B enhancement. Furthermore, the finding that saving List B did not enhance recall for List A items also supports the proactive interference account as presenting List B after List A would have interfered with List A recall given that the save cue was only presented to participants after encoding List B. Taken together, the results of the first experiment suggest that these two accounts, although theoretically different, might not be mutually exclusive in producing the saving-enhanced memory effect.

In terms of memory performance for the lists that were saved, Experiment 1 found that the proportion of words recalled from a saved list was higher than the proportion of words recalled from an unsaved list. This finding supports that of Storm and Stone (2015) who found better recall for List A items when List A was saved compared to when it was not saved. This finding is also in line with previous research showing that accuracy is better when participants offload information (see Gilbert et al., 2022 for a review). In accordance with this, participants in Experiment 2 not only showed a preference towards saving but also showed a preference towards saving List A when List B was tested first. Given the results of Experiment 1, saving List A when List B is tested first seems to be most beneficial for both List A recall and List B recall.

Previous research has also shown that offloading is associated one's confidence in their memory abilities (e.g., Boldt & Gilbert, 2019; Gilbert, 2015b; Gilbert et al., 2020; Hu et al., 2019). However, Experiment 2 failed to replicate the association between pre-task or post-task confidence and decision to offload. This being said, participants' decision on which list to offload was correlated with difference between their post-task confidence ratings for

List B recall when List A was saved post-task confidence ratings for List A recall when List B was saved. This finding suggests that individuals have some insight on their memory performance and that they accordingly monitor their confidence in recalling the unsaved list. Since this association between confidence and choice of saving was only found with post-task confidence, it is likely that post-task confidence judgements were revised during the task and so were more associated with participant's offloading strategy than pre-task confidence judgements (Boldt et al., 2019). Further exploring the role of metacognition in this line of research could help individuals improve their use of offloading strategies to gain more benefits.

Conclusion

In conclusion the current chapter adds to our understanding of the consequences of offloading on subsequently presented information. The first experiment found that saving the first list and recalling the second list leads to enhanced memory for both the saved and unsaved lists. The second experiment found that when given a choice, participants not only choose to save a list but also choose to save the first list (List A) in the condition where it is recalled second (i.e., List B is tested first). So, when given a choice, participants tend to choose the list that would be more beneficial.

CHAPTER 6. GENERAL DISCUSSION

Overview

The research presented in this thesis investigated the causes and consequences of cognitive offloading. In particular, it explored the role of confidence and effort-minimization in intention offloading, and the consequences of offloading on subsequent items.

Chapter 2 investigated whether metacognitive interventions designed to influence participants' confidence subsequently influenced reminder bias in an intention offloading task. Chapter 3 tested the role of effort-minimization in intention offloading. Chapter 4 examined domain-general confidence signals in cognitive offloading where this chapter investigated whether confidence in a task from an unrelated domain is related to reminder bias in an intention offloading task. Finally, chapter 5 presented two experiments where the first experiment tested the consequences of cognitive offloading on memory for subsequent items and the second experiment investigated A) whether participants have a preference towards saving material, and B) if they do, which material they choose to offload.

This final chapter will briefly summarize each piece of research. Then, I will discuss how the separate findings come together and contribute to a wider understanding of cognitive offloading. Finally, the limitations of the present research and future discussions will be discussed.

Research summaries

The role of confidence in intention offloading

The experiment presented in chapter 2 investigated the role of metacognition in intention offloading. This experiment adds to the growing body of research investigating how and when individuals decide to use external resources to support their memory for delayed intentions (e.g., Ball et al., 2021; Gilbert, 2015a, 2015b; Gilbert et al., 2020; Kirk et al., 2021). One finding from this line of research is that individuals' confidence in their memory abilities influences their decision to use reminders (Gilbert, 2015b; Gilbert et al., 2020 Experiment 2; Kirk et al., 2021). Following on from this finding, the experiment presented in this chapter explored whether metacognitive interventions designed to influence participants' confidence in their internal memory abilities would in turn influence their decision to use external reminders. A finding such as this would substantiate the role of metacognition in cognitive offloading.

To evaluate the above hypothesis, the study presented in this chapter adapted the paradigm developed by Gilbert et al. (2020 Experiment 2) where participants were presented with two metacognitive interventions, feedback valence and practice difficulty. For the first intervention, half of the participants received positive feedback on their performance while the other half received negative feedback. For the second intervention, half of the participants received four easy practice trials where 4 out of 25 circles were targets while the other half received four difficult practice trials where 16 out of 25 circles were targets. Both these interventions were manipulated in a 2 x 2 between-subjects design yielding 4 groups, easy-practice/positive-feedback, easy-practice/negative-feedback, difficult-practice/positive-feedback and difficult-practice/negative-feedback.

The results of this experiment showed that both interventions influenced participants' metacognitive bias without influencing their accuracy. So even though participants' confidence differed in the four groups, their accuracy did not. These metacognitive interventions also had a parallel effect on participants' bias towards reminders where participants in the difficult-practice/negative-feedback group were the most underconfident and biased towards using reminders. It was also found that these shifts in reminder bias were mediated by shifts in participants confidence.

Although evidence for a parallel effect on reminder bias was found, the results also demonstrated that this reminder bias cannot be fully explained by confidence as participants in the easy-positive/positive-feedback group were simultaneously overconfident and biased towards using reminders. Therefore, the bias towards reminders was seen in conjunction with both over- and under-confidence. Thus, it was concluded that metacognition is not the only factor influencing reminder bias and that there must be additional factors contributing to this bias.

The role of effort-minimization in cognitive offloading

The aim of chapter 3 was to investigate an additional factor that might contribute to offloading behaviour. This chapter proposed that one additional factor might be preference to avoid cognitive effort associated with remembering intentions internally (i.e., effort-minimization).

To investigate this, the paradigm utilized in chapter 2 was modified to include performance-based rewards. Because chapter 2 found that participants in the easy-practice/positive-feedback group were simultaneously overconfident and biased towards using reminders, only this condition was replicated in the experiment presented in this chapter. One group of participants received monetary compensation contingent on their performance in the task (reward group) while the other group received a flat payment for taking part in the experiment (no-reward group). It was hypothesized that to earn more money, participants in the reward group would increase their effort in performing the task by using their own memory, thus reducing (but not necessarily eliminating) bias towards setting reminders.

Results from this experiment found that bias towards using reminders was significantly reduced in the group that received performance-based rewards. This suggests that preference to avoid cognitive effort influences cognitive behaviour such that individuals are more willing to allocate cognitive effort when monetary rewards are attached to a task.

A secondary aim of this chapter was to investigate whether participants might become increasingly underconfident with practice (i.e., the UWP effect) (see Koriat et al., 2002) as this could also explain a bias towards external reminders. This was measured by adding an additional confidence judgement at the end of the experiment and comparing the difference between the ratings given on that judgement and the ratings provided on the confidence judgement at the start of the main task (after the practice trials). The results of this experiment did not find support for the UWP effect. Although participants were less overconfident at the end of the experiment, their confidence ratings did not differ significantly between their initial and their final ratings.

Domain-general versus task-specific metacognitive signals in cognitive offloading

Chapter 2 found a relationship between confidence and intention offloading where participants' confidence was able to account, at least in part, for their bias towards using reminders. An association between metacognition and strategy choice has also been found in other domains (see Gilbert, 2015b; Gilbert et al., 2020 in strategic intention offloading; see Nelson & Narens, 1990 in memory; see Yeung & Summerfield, 2012 in decision-making). Since metacognition has been found to influence strategy choice in various domains, the aim of chapter 4 was to examine the extent to which intention offloading is linked to domain-general versus task-specific confidence signals.

To investigate this, the study presented in this chapter employed three tasks, one memory task (adapted version of the intention offloading task presented in chapter 2) and two perceptual tasks adapted from Gilbert (2015b Experiment 2). In the two perceptual tasks participants' objective accuracy was equalized using a staircase procedure. Also in the perceptual tasks, two confidence measures were derived. The first was a measure of metacognitive bias (calculated as mean confidence across trials) and the second was a measure of metacognitive efficiency. In the intention offloading task, two offloading measures were collected. The first was the likelihood of setting reminders (i.e., propensity to offload) and the second was bias in reminder-setting behaviour (i.e., preference to offload, relative to the optimal strategy).

The results of this experiment found evidence for the influence of both domain-general and task-specific confidence signals on reminder-setting behaviour where perceptual confidence was related to the first (i.e., propensity to offload) but not the second (i.e., preference to offload) intention offloading measure. There was no evidence for a link between perceptual metacognitive efficiency and reminder-setting behaviour.

Consequences of cognitive offloading

Chapters 2 to 4 investigated factors influencing intention offloading. Chapter 5 examined how cognitive offloading might affect memory for subsequent information. Research in the domain of cognitive offloading has found that saving one computer file before studying a new file significantly improves recall of the contents of not only the saved file but also the contents of the new file (Runge et al., 2019, 2020; Storm & Stone, 2015). In contrast findings from directed forgetting research have been equivocal where some studies have found that forgetting the middle list led to reduced recall not only for this list but also for the first list which was not intended for forgetting (Racsmány et al., 2019; Sahakyan, 2004) while other studies have found that participants are able to selectively forget the items from the first list without forgetting items from the second list (Kliegl et al., 2013, 2020a).

One difference between studies conducted in the domains of cognitive offloading and directed forgetting is that in cognitive offloading research participants always had to save the first list while in directed forgetting research participants are also cued to forget the second list on some trials. In the domain of directed forgetting, there is one discrepancy where in Sahakyan's (2004) experiment, participants had to recall List A before List B while in Kliegl et al.'s (2013) experiment, this order was counterbalanced between participants.

The study presented in this chapter consisted of two experiments that aimed to extend on previous research (or perhaps even reconcile) by investigating, 1) whether the saving-enhanced memory effect holds when the to-be-remembered information is presented before the saved information and 2) people's offloading preferences when given two pieces of information to remember. In both experiments recall order of the two lists was manipulated where half of the participants had to recall List A before recalling List B while for the other half this order reversed.

Experiment 1 also manipulated whether and which list was saved where participants could either save List A, List B or not save either list. It was found that saving a list always resulted in enhanced memory for that list. Furthermore, saving List A led to a memory enhancement for List B only when List B was recalled first. But saving List B did not enhance recall for List A regardless of List A's testing order. These findings suggest that, in terms of saving-enhanced memory, saving List A might prove most beneficial.

In experiment 2, participants were allowed to choose their preferred strategy. It was found that participants chose to save List A significantly more often than List B but only when List B was tested first. This corresponds with the findings of experiment 1 where saving List A resulted in a memory enhancement for List B only when List B was tested first.

Bringing it all together

All experiments presented in this thesis show that when given a choice between choosing to offload versus using unaided memory, participants prefer to offload (as shown by a bias towards reminders in chapters 2 to 4 and a preference towards saving a list in chapter 5).

Although research has found multiple factors influencing intention offloading (see Gilbert et al., 2022 for a review), the current thesis focused on exploring the role of metacognition and effort-minimization in intention offloading. Previous research has established the role of metacognition in intention offloading (e.g., Boldt & Gilbert, 2019; Gilbert, 2015b; Gilbert et al., 2020). The results of the second chapter extended on previous research by finding that metacognitive interventions designed to shift people's confidence also shifted their bias towards reminders. This finding has important implications as this means that metacognitive interventions such as the ones outlined in chapter 2 could play an important role in optimizing individuals' use of reminders. As individuals rely on offloading to organize their daily life but do not always use external resources optimally, finding interventions to influence individuals' offloading strategies could improve this organization in everyday life. These interventions would be especially useful for older adults who have physical or neurological impairments that affect independent living as these interventions can improve independence and ability to complete everyday tasks (Berry et al., 2010).

Previous research investigating the effect of metacognitive intervention training on participants' metacognitive evaluations have found metacognitive interventions to have a strong impact on both, task-specific (e.g., Engeler & Gilbert, 2020; Vorwerk et al., 2022a) and domain-general (e.g., Carpenter et al., 2019) metacognition. Research has identified various metacognitive interventions that could be useful when training individual's confidence. For example, Engeler and Gilbert (2020) and Carpenter et al. (2019) asked participants to provide performance predictions after which the experimenters provided feedback to participants on the accuracy of their performance judgements. Vorwerk et al. (2022a) used cognitive-based interventions and taught memory strategies to older adults. Older adults then had a chance to practice these strategies in everyday life. Vorwerk et al. (2022a) found that this led to more optimistic beliefs regarding participants' memory abilities. Furthermore, they also found that these cognitive-based interventions led to a greater knowledge on memory strategies that can be used to supplement their memory.

However, it should be noted that there are individual differences in metacognitive bias where some studies have found overconfidence in participants in terms of their memory abilities (e.g., Cauvin et al., 2019; Knight et al., 2005; Schnitzspahn et al., 2011) and others have found underconfidence in participants with regards to memory abilities (Gilbert, 2015b; Gilbert et al., 2020). This means that metacognitive interventions that can potentially remedy biases in either direction, taking individual differences into account, should be created.

Although previous research has found metacognitive interventions to be effective at calibrating individuals' confidence, their impact has not necessarily translated to offloading behaviour (see Engeler & Gilbert, 2020; Grinschgl et al., 2020). For example, Engeler and Gilbert (2020) did not find an influence of metacognitive interventions on reminder bias. Furthermore, Grinschgl et al. (2020) found that fake performance feedback influenced participants' metacognitive evaluations of their cognitive ability but did not influence offloading behaviour in a working memory task (although participants who received below-average feedback reported relying more on offloading strategies). Although the results of previous research is mixed in terms of metacognitive intervention training influencing offloading behaviour, it is not to say that metacognitive intervention training would not translate to offloading behaviour. The results of chapter 2 still show that metacognitive interventions can shift participants' reminder bias. So, it may well be that appropriate interventions need to be designed to see an impact of metacognitive interventions on offloading behaviour. For example, a more real world cognitive-based intervention could be developed to examine shifts in reminder bias (as in Vorwerk et al., 2022a). Furthermore, randomized controlled trials using gamification as interventions have found (although mixed) positive shifts in participants' behaviours in the domains of health and education (Beemer et al., 2019; see Johnson et al., 2016 for a systematic review; Patel et al., 2019). Future research could explore whether gamification interventions have an impact on participants' reminder-setting behaviour.

Metacognitive intervention training could be a very desirable avenue to optimise individuals' strategy choice. The reason for this being the domain-general component of metacognition. This was found in chapter 4 of this thesis where perceptual confidence correlated not only with memory confidence but also with individuals' propensity to offload. This suggests that training individuals' confidence in one domain could shift their confidence in multiple task domains including that of memory, ultimately influencing their strategy choice. This potentially makes metacognitive training a cost-effective intervention to optimise individuals' behaviours. Reminder bias (i.e., preference to offload) however, was found to have been influenced by task-specific confidence signals. Therefore, it becomes important to disentangle propensity to offload from preference to offload and understand factors that influence both these measures. This is so that more effective interventions can be designed to optimise offloading.

The results of chapter 2 also suggest that metacognition is not the only factor influencing reminder bias. This also supports the findings of Engeler and Gilbert (2020) who found that although participants who received feedback on their performance judgements were well-calibrated in their confidence judgements, they were still biased towards using reminders. Accordingly, chapter 3 found that effort-minimization is another factor that influences reminder bias. Although the results of this study could suggest that individuals have an intrinsic drive to avoid using internal memory (see Ballard et al., 1997a; Hull, 1943), the results also support a different account proposed by Kurzban et al. (2013). This account suggests that effortful activities engage domain-general cognitive processes that are limited in capacity. So, to redirect these processes towards other tasks, individuals avoid expending cognitive effort. In the experiment presented in chapter 3, perhaps the reason why participants were still biased towards using reminders was simply because they were trying to conserve their already limited cognitive processes.

Further support for this theory comes from the first experiment presented in chapter 5 where saving a list resulted in better memory for subsequent information (also see Runge et al., 2019, 2020; Storm & Stone, 2015). This means that when a list was saved, participants were able to redirect their cognitive resources to studying subsequent information. However, this was only true when the to-be-remembered information (i.e., List B) was presented after the saved information (i.e., List A). When the to-be-remembered information (i.e., List A) was presented before the saved information (i.e., List B), there was no benefit to participants’ recall accuracy. As mentioned in chapter 5, this particular finding supports the account of proactive interference where presenting List B after List A induced proactive interference where both lists were perhaps competing for cognitive resources. In line with this the second experiment presented in chapter 5 found that participants had a preference towards saving List A only when List B was tested first. From experiment 1, we know that this was the most beneficial strategy as saving List A in this case enhanced recall performance for both List A and List B.

Taken together, the findings of all experiments show the ubiquitous nature of cognitive offloading. What they also show is that cognitive offloading is influenced by multiple factors which despite being theoretically different, might not be mutually exclusive when it comes to influencing offloading. Furthermore, from the findings of this thesis, it is clear that cognitive offloading improves task performance and if used optimally, it can be a powerful tool to guide organisation of behaviour in everyday life.

Future directions

An obvious limitation to the studies described above is that they are all laboratory-based experiments. Therefore, we do not know how offloading strategies in the laboratory might generalize to offloading strategies in the real world. Future research should draw parallels between the two where they could potentially explore whether factors influencing offloading strategies in the laboratory also correlate with factors influencing offloading strategies in the real-world. This would be especially useful in closing the gap in ageing research where findings have been inconsistent between laboratory and naturalistic settings (e.g., Cauvin et al., 2019; Schnitzspahn et al., 2011).

A second avenue that should be explored by future research are the various factors influencing cognitive offloading behaviour. Although metacognition seems to be a well-established factor influencing cognitive offloading (e.g., Boldt & Gilbert, 2019; Dunn & Risko, 2016; Gilbert, 2015b; Gilbert et al., 2020; Hu et al., 2019), there are also studies that did not find an influence of metacognition on cognitive offloading (Grinschgl et al., 2020). Future research should investigate conditions under which factors are shown to influence cognitive offloading. For example, Grinschgl et al. (2020) differentiate between metacognitive beliefs and metacognitive experiences when explaining these diverging results, where metacognitive beliefs refer to general beliefs about one's memory abilities and metacognitive experiences refer to task-specific knowledge. Perhaps future research could manipulate this between-subjects within a single experiment to elucidate conditions under which offloading might be influenced by metacognition.

A third avenue that should be pursued by future research is the domain-general component of cognitive offloading. Gilbert (2015b) and the study presented in chapter 4 found that individuals' propensity to offload was correlated with their memory confidence and their perceptual confidence suggesting that propensity to offload is associated with domain-general metacognitive signals. However, chapter 4 also found that individuals' preference to offload was not associated with domain-general confidence signals but was influenced by task-specific confidence signals. Furthermore, results have been inconsistent when exploring the relationship between cognitive offloading in different domains. While Ball et al. (2021) found that intention offloading strategies correlate across three versions of the same task, Meyerhoff et al. (2021) did not find significant associations between cognitive offloading strategies measured in an intention offloading task and offloading strategies in a perceptual block-copy task. As we know very little about this, future research should explore the task-specific versus domain-general contributions to cognitive offloading strategies. Future research should also investigate the impact of metacognitive training on these strategies. This will aid in the development of real-world metacognitive interventions to improve individuals' cognitive offloading strategies. As mentioned in the previous section, results from the avenue of metacognitive interventions is still mixed, so future research should help progress the issue to find interventions that would translate to everyday behavioural strategies.

The fourth avenue that could be exmained by future research is cognitive offloading in the synaesthetic population. For example, research has found enhanced performance by grapheme-colour synaesthetes on visual memory tasks (e.g., Rothen & Meier, 2010; Yaro & Ward, 2007). It would be interesting to see whether this memory advantage extends to memory for delayed intentions and how this would translate to intention offloading. For example, how do offloading strategies differ between synaesthetes and controls and does this translate to better accuracy in a delayed intentions task. The reason for this is because research has also found that adults can be trained to acquire synaesthetic experiences (e.g., Bor et al., 2014). Perhaps if future research finds enhanced accuracy in intention fulfilment in synaesthetes, this would provide an avenue for intervention training in non-synaesthetes.

The fifth avenue that should be pursued by future research is the investigation of offloading strategies in psychological disorders such as obsessive-compulsive disorder (OCD). The case of OCD is particularly interesting because previous research has found a dissociation in individuals with OCD where they are able to update their confidence just as accurately as neurotypicals but they fail to use it to guide behaviour (Vaghi et al., 2017). Seeing as research has found an association between confidence and cognitive offloading, it would be interesting to investigate this link in individuals with OCD and potentially other psychological disorders.

Lastly, research should elucidate factors involved in saving one list over the other. Clearly, the saving-enhanced memory effect can only be seen under certain circumstances. Therefore, it would be beneficial if research, 1) investigated more circumstances under which this was observed, and 2) investigated factors contributing to the decision to save one list over the other. For example, future research could explore saving-enhanced memory in three list paradigms, i.e., what happens when individuals are presented with more information that they have to study and subsequently recall. Furthermore, future research should also look at the role of metacognition in list-saving behaviour. Although our study did not find an association between pre-task confidence and list-saving, this result should be interpreted with caution because, 1) it rests on a null finding and, 2) the study was not powered to detect an effect of confidence on list-saving. The study did, however, find an effect of post-task confidence on List A/List B saving. Therefore, future research should further investigate the role of both pre-task and post-task confidence judgements on list-saving behaviour.

Conclusions

With the rapid advancement of technology in today's world, cognitive offloading has become increasingly ubiquitous. The work conducted in this thesis aimed to understand the causes and consequences of cognitive offloading. The reported findings suggest that there are likely multiple factors that contribute to the phenomenon of cognitive offloading. These factors include metacognition and effort-minimization. With regards to metacognition, it was found that participants who were more underconfident in their memory abilities tended to set more reminders and that interventions designed to shift participants' confidence subsequently shifted their bias towards reminders. It was also found that propensity to offload was influenced by domain-general metacognitive signals. With regards to effort-minimization it was found that administering performance-based financial rewards reduced participants' reminder bias. Moreover, it seems that offloading information frees cognitive resources to help remember subsequent information but that this is dependent on which list is saved and in what order the information is recalled. Together, the findings of this thesis contribute to research in cognitive offloading. Yet still, research in cognitive offloading is in its infancy and there are many open questions waiting to be explored.

References

Aarts, E., Roelofs, A., Franke, B., Rijpkema, M., Fernández, G., Helmich, R. C., & Cools, R. (2010). Striatal dopamine mediates the interface between motivational and cognitive control in humans: Evidence from genetic imaging. *Neuropsychopharmacology*, *35*(9), 1943–1951. https://doi.org/10.1038/npp.2010.68

Aguirre, C., Gómez-Ariza, C. J., & Bajo, M. T. (2020). Selective directed forgetting: Eliminating output order and demand characteristics explanations. *Quarterly Journal of Experimental Psychology*, *73*(9), 1514–1522. https://doi.org/10.1177/1747021820915100

Baird, B., Cieslak, M., Smallwood, J., Grafton, S., & Schooler, J. (2015). Regional white matter variation associated with domain-specific metacognitive accuracy. *Journal of Cognitive Neuroscience*, *27*(3), 440–452. https://doi.org/https://doi.org/10.1162/jocn_a_00741

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

Baldwin, V. N., Powell, T., & Lorenc, L. (2011). Factors influencing the uptake of memory compensations: A qualitative analysis. *Neuropsychological Rehabilitation*, *21*(4), 484–501. https://doi.org/https://doi.org/10.1080/09602011.2011.582378

Ball, H. B., & Brewer, G. A. (2018). Proactive control processes in event-based prospective memory: Evidence from intraindividual variability and ex-gaussian analyses. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*(5), 793–811. https://doi.org/10.1037/xlm0000489

Ball, H. B., Knight, J. B., Dewitt, M. R., & Brewer, G. A. (2013). Individual differences in the delayed execution of prospective memories. *Quarterly Journal of Experimental Psychology*, *66*(12), 2411–2425. https://doi.org/10.1080/17470218.2013.785576

Ball, H. B., Peper, P., Alakbarova, D., Gilbert, S., & Brewer, G. A. (2021). Individual differences in working memory capacity predict benefits to memory from intention offloading. *Memory*, *21*(1), 1–15. https://doi.org/10.1080/09658211.2021.1991380

Ballard, D. H., Hayhoe, M. M., Li, F., & Whitehead, S. (1992). Hand-eye coordination during sequential tasks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *337*, 331–339. https://doi.org/10.1098/rstb.1992.0111 76948680.1177/0956797617694868

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*(1), 66–80. https://doi.org/10.1162/jocn.1995.7.1.66

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997a). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*(4), 723–742. http://journals.cambridge.org/abstract_S0140525X97001611

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997b). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*(4), 723–767. https://doi.org/10.1017/S0140525X97001611

Bäuml, K. H., Pastötter, B., & Hanslmayr, S. (2010). Binding and inhibition in episodic memory-Cognitive, emotional, and neural processes. In *Neuroscience and Biobehavioral Reviews* (Vol. 34, Issue 7, pp. 1047–1054). Pergamon. https://doi.org/10.1016/j.neubiorev.2009.04.005

Beemer, L. R., Ajibewa, T. A., Dellavecchia, G., & Hasson, R. E. (2019). A pilot intervention using gamification to enhance student participation in classroom activity breaks.

International Journal of Environmental Research and Public Health, *16*(21), 1–11. https://doi.org/10.3390/ijerph16214082

Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiology and Behavior*, *81*(2), 179–209. https://doi.org/10.1016/j.physbeh.2004.02.004

Berry, J. M., Hastings, E. C., West, R. L., Courney, L., & Cavanaugh, J. C. (2010). Memory aging: Deficits, beliefs, and interventions. In J. C. Cavanaugh, C. K. Cavanaugh, J. M. Berry, & R. L. West (Eds.), *Aging in America. Psychological Aspects*.

Bilalić, M., McLeod, P., & Gobet, F. (2008). Inflexibility of experts - reality or myth? Quantifying the Einstellung effect in chess masters. *Cognitive Psychology*, *56*(2), 73–102. https://doi.org/10.1016/J. COGPSYCH.2007.02.001

Boldt, A., & Gilbert, S. J. (2019). Confidence guides spontaneous cognitive offloading. *Cognitive Research: Principles and Implications*, 1–33. https://doi.org/10.31234/osf.io/ct52k

Boldt, A., Schiffer, A. M., Waszak, F., & Yeung, N. (2019). Confidence predictions affect performance confidence and neural preparation in perceptual decision making. *Scientific Reports*, *9*(1), 1–17. https://doi.org/10.1038/s41598-019-40681-9

Bor, D., Rothen, N., Schwartzman, D. J., Clayton, S., & Seth, A. K. (2014). Adults can be trained to acquire synesthetic experiences. *Scientific Reports*, *4*. https://doi.org/10.1038/srep07089

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology*, *66*, 83–113. https://doi.org/https://doi.org/10.1146/annurev-psych-010814-015044

Brandimonte, Einstein, G. O., & McDaniel, M. A. (2014). Prospective memory: Theory and applications. In *Prospective Memory*. Psychology Press.

https://doi.org/10.4324/9781315806488

Brügger, A., Richter, K. F., & Fabrikant, S. I. (2019). How does navigation system behavior influence human behavior? *Cognitive Research: Principles and Implications*, *4*(1). https://doi.org/10.1186/s41235-019-0156-5

Carpenter, J., Sherman, M., Kievit, R., Seth, A., Lau, H., & Fleming, S. (2019). Domain-general enhancements of metacognitive ability through adaptive training. *Journal of Experimental Psychology: General*, *148*(1), 51–64. https://doi.org/http://dx.doi.org/10.1037/xge0000505

Cauvin, S., Moulin, C., Souchay, C., Schnitzspahn, K., & Kliegel, M. (2019). Laboratory vs. naturalistic prospective memory task predictions: Young adults are overconfident outside of the laboratory. *Memory*, *27*(5), 592–602. https://doi.org/10.1080/09658211.2018.1540703

Clark, A. (2010). *Supersizing the mind*. Oxford University Press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7–19.

Cockburn, J., & Smith, P. T. (1988). Effects of age and intelligence on everyday memory tasks. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues. Volume 2: Clinical and educational implications* (pp. 132–136). Wiley.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*(1), 51–57. https://doi.org/10.1177/0963721409359277

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

Dreisbach, G., & Fischer, R. (2015). Conflicts as aversive signals for control adaptation. *Current Directions in Psychological Science*, *24*(4), 255–260. https://doi.org/10.1177/0963721415569569

Dunn, T. L., Lutes, D. J. C., & Risko, E. F. (2016). Metacognitive evaluation in the avoidance of demand. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1372–1387. https://doi.org/https://doi.org/10.1037/xhp0000236

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*, Advance online publication. 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. https://doi.org/10.1037/0278-7393.16.4.717

Eisenberg, R. (1992). Learned industriousness. *Psychological Review*, *99*(2), 248–267.

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

Engeler, N. C., & Gilbert, S. J. (2020). The effect of metacognitive training on confidence and strategic reminder setting. *PLoS ONE*, *15*(10 October 2020), 1–14. https://doi.org/10.1371/journal.pone.0240858

Faivre, N., Filevich, E., Solovey, G., Kühn, S., & Blanke, O. (2017). Behavioural, modeling, and electrophysiological evidence for domain-generality in human metacognition. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *Advance* *on*. https://doi.org/http://dx.doi.org/10.1523/jneurosci.0322-17.2017

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fish, J., Wilson, B. A., & Manly, T. (2010). The assessment and rehabilitation of prospective memory problems in people with neurological disorders: A review. *Neuropsychological Rehabilitation*, *20*(2), 161–179. https://doi.org/10.1080/09602010903126029

Fleming, J., Ownsworth, T., Doig, E., Hutton, L., Griffin, J., Kendall, M., & Shum, D. H. K. (2017). The efficacy of prospective memory rehabilitation plus metacognitive skills training for adults with traumatic brain injury: Study protocol for a randomized controlled trial. *Trials*, *18*(1), 1–11. https://doi.org/10.1186/s13063-016-1758-6

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*(July), 1–9. https://doi.org/10.3389/fnhum.2014.00443

Fröber, K., & Dreisbach, G. (2014). The differential influences of positive affect, random reward, and performance-contingent reward on cognitive control. *Cognitive, Affective and Behavioral Neuroscience*, *14*(2), 530–547. https://doi.org/10.3758/s13415-014-0259-x

Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin and Review*, *10*(4), 843–876. https://doi.org/10.3758/BF03196546

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. M., 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., Bird, G., Frith, C. D., & Burgess, P. W. (2012). Does “task difficulty” explain “task-induced deactivation?” *Frontiers in Psychology*, *3*(125), 1–12. https://doi.org/10.3389/fpsyg.2012.00125

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 and Review*, *0123456789*. https://doi.org/10.3758/s13423-022-02139-4

Gilchrist, A. L., Cowan, N., & Naveh-Benjamin, M. (2008). Working memory capacity for spoken sentences decreases with adult ageing: Recall of fewer but not smaller chunks in older adults. *Memory*, *16*(7), 773–787. https://doi.org/10.1080/09658210802261124

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*(6), 516–522. https://doi.org/10.1111/1467-9280.00395

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., Meyerhoff, H. S., Schwan, S., & Papenmeier, F. (2020). From metacognitive beliefs to strategy selection: Does fake performance feedback influence cognitive offloading? *Psychological Research*, *0123456789*. https://doi.org/10.1007/s00426-020-01435-9

Guo, Y., Gan, J., Ping, Y., Song, T., Liu, T., Wang, E., & Li, Y. (2021). Effective external reminders impair the practice effect of time-based prospective memory. *International Journal of Psychology*. https://doi.org/10.1002/ijop.12821

Guynn, M. J., McDaniel, M. A., & Einstein, G. O. (1998). Prospective memory: When reminders fail. *Memory and Cognition*, *26*(2), 287–298. https://doi.org/10.3758/BF03201140

Hayes, A. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. https://books.google.co.uk/books?hl=en&lr=&id=8ZM6DwAAQBAJ&oi=fnd&pg=PP1&dq=introduction+to+mediation+moderation+and+conditional+process+analysis+second+edition&ots=20DdqPZk-B&sig=oVM-8maNtTh3O1f9maeU2l0U9hA

Heed, T., Habets, B., Sebanz, N., & Knoblich, G. (2010). Others' actions reduce crossmodal integration in peripersonal space. *Current Biology*, *20*(15), 1345–1349. https://doi.org/10.1016/j.cub.2010.05.068

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

Holbrook, J., & Dismukes, K. (2009). Prospective memory in everyday tasks. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting*, 590–594. https://doi.org/10.4324/9781315749860-21

Hu, X., Luo, L., & Fleming, S. M. (2019). A role for metamemory in cognitive offloading. *Cognition*, *193*, 104012. https://doi.org/10.1016/j.cognition.2019.104012

Hull. (1943). *Principles of behavior*. Appleton Century-Croft.

Hutchins, E. (1995). *Cognition in the wild*. MIT Press.

Inzlicht, M., Shenhav, 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

Jimura, K., Locke, H. S., & Braver, T. S. (2010). Prefrontal cortex mediation of cognitive enhancement in rewarding motivational contexts. *Proceedings of the National Academy of Sciences*, *107*(19), 8871–8876. https://doi.org/10.1073/pnas.1002007107

Johnson, D., Deterding, S., Kuhn, K. A., Staneva, A., Stoyanov, S., & Hides, L. (2016). Gamification for health and wellbeing: A systematic review of the literature. *Internet Interventions*, *6*, 89–106. https://doi.org/10.1016/j.invent.2016.10.002

Kelly, M. O., & Risko, E. F. (2019). Offloading memory: Serial position effects. *Psychonomic Bulletin and Review*, *26*(4), 1347–1353. https://doi.org/10.3758/s13423-019-01615-8

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

Kirsh, D. (1996). Adapting the environment instead of oneself. *Adaptive Behavior*, *4*, 415–452. https://doi.org/10.1177/105971239600400307

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*(4), 513–549. https://doi.org/10.1016/0364-0213(94)90007-8

Kliegel, M., & Martin, M. (2003). Prospective memory research: Why is it relevant? *International Journal of Psychology*, *38*(4), 193–194. https://doi.org/10.1080/00207590344000114

Kliegl, O., Pastötter, B., & Bäuml, K.-H. T. (2020a). Does amount of pre-cue encoding modulate selective list method directed forgetting? *Frontiers in Psychology*, *11*, 1403. https://doi.org/10.3389/fpsyg.2020.01403

Kliegl, 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 and Cognition*, *41*(3), 452–464. https://doi.org/10.3758/s13421-012-0264-7

Kliegl, O., Pastötter, B., & Bäuml, K. H. T. (2020b). Does Amount of Pre-cue Encoding Modulate Selective List Method Directed Forgetting? *Frontiers in Psychology*, *11*(July). 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/10.2996/kmj/1138846322

Koriat, A. (2007). *Metacognition and consciousness*. Cambridge University Press.

Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General*, *131*(2), 147–162. https://doi.org/10.1037/0096-3445.131.2.147

Krebs, R. M., Boehler, C. N., & Woldorff, M. G. (2010). The influence of reward associations on conflict processing in the stroop task. *Cognition*, *117*(3), 341–347. https://doi.org/10.1016/j.cognition.2010.08.018

Kurzban, R. (2016). The sense of effort. *Current Opinion in Psychology*, *7*, 67–70.

https://doi.org/10.1016/j.copsyc.2015.08.003

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(6), 661–679. https://doi.org/10.1017/S0140525X12003196

Lee, A. L. F., Ruby, E., Giles, N., & Lau, H. (2018). Cross-domain association in metacognitive efficiency depends on first-order task types. *Frontiers in Psychology*, *9*, 2462. https://doi.org/10.3389/fpsyg.2018.02464

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/https://doi.org/10.1016/j.jesp.2013.03.013

Leys, Christophe, Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Locke, H. S., & Braver, T. S. (2008). Motivational influences on cognitive control: Behavior, brain activation, and individual differences. *Cognitive, Affective and Behavioral Neuroscience*, *8*(1), 99–112. https://doi.org/10.3758/CABN.8.1.99

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*(1), 422–430. https://doi.org/10.1016/j.concog.2011.09.021

Mäntylä, T. (1996). Activating actions and interrupting intentions: Mechanisms ofretrieval sensitization in prospective memory. In M. A. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 93–113). Erlbaum.

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/j.jarmac.2018.11.001

Marsh, R. L., & Hicks, J. L. (1998). Event-based prospective memory and executive control of working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 336–349. https://doi.org/10.1037/0278-7393.24.2.336

Marsh, R. L., Hicks, J. L., & Landau, J. D. (1998). An investigation of everyday prospective memory. *Memory and Cognition*, *26*(4), 633–643. https://doi.org/10.3758/BF03211383

Mazancieux, A., Fleming, S. M., Souchay, C., & Moulin, C. J. A. (2020). Is There a G Factor for Metacognition? Correlations in Retrospective Metacognitive Sensitivity Across Tasks. *Journal of Experimental Psychology: General*, *149*(9), 1788–1799. https://doi.org/10.1037/xge0000746.supp

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

McWilliams, A., Bibby, H., Steinbeis, N., David, A. S., & Fleming, S. M. (2022). Age-related decreases in global metacognition are independent of local metacognition and task performance. *Psyarxiv*, 1–23. https://doi.org/10.31234/osf.io/nmhxv

Meacham, J. A., & Leiman, B. (1982). Remembering to perform future actions. In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 327–336). Freeman.

Menary, R. (2010). No Title. In *The extended mind*. MIT Press.

Meyerhoff, H. S., Grinschgl, S., Papenmeier, F., & Gilbert, S. J. (2021). Individual differences in cognitive offloading: A comparison of intention offloading, pattern copy, and short-term memory capacity. *Cognitive Research: Principles and Implications*, *6*(34). https://doi.org/10.1186/s41235-021-00298-x

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *Journal of Neuroscience*, *38*(14), 3534–3546. https://doi.org/10.1523/JNEUROSCI.2360-17.2018

Morita, T. (2006). Reminders supporting spontaneous remembering in prospective memory tasks. *Japanese Psychological Research*, *48*(1), 34–39. https://doi.org/10.1111/j.1468-5884.2006.00303.x

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation*. New York, NY: Academic Press. https://doi.org/10.1016/S0079-7421(08)60053-5

Nestojko, J. F., Finley, J. R., & Roediger, H. L. (2013). Extending cognition to external agents. *Psychological Inquiry*, *24*(4), 321–325. https://doi.org/10.1080/1047840X.2013.844056

Niv, Y., Joel, D., & Dayan, P. (2006). A normative perspective on motivation. *Trends in Cognitive Sciences*, *10*(8), 375–381. https://doi.org/10.1016/j.tics.2006.06.010

Norman, D. (1993). Cognition in the head and in the world : An introduction to the special issue on situated action. *Cognitive Science*, *17*, 1–6.

Otto, A. R., & Vassena, E. (2020). It's all relative: Reward-induced cognitive control modulation depends on context. *Journal of Experimental Psychology: General*.

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

Pansky, A., & Goldsmith, M. (2014). Metacognitive effects of initial question difficulty on subsequent memory performance. *Psychonomic Bulletin and Review*, *21*(5), 1255–1262. https://doi.org/10.3758/s13423-014-0597-2

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, *52*(3), 381–410. https://doi.org/10.1177/0018720810376055.

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 and Cognition*, *40*(6), 861–873. https://doi.org/10.3758/s13421-012-0206-4

Patel, M. S., Small, D. S., Harrison, J. D., Fortunato, M. P., Oon, A. L., Rareshide, C. A. L., Reh, G., Szwartz, G., Guszcza, J., Steier, D., Kalra, P., & Hilbert, V. (2019). Effectiveness of Behaviorally Designed Gamification Interventions with Social Incentives for Increasing Physical Activity among Overweight and Obese Adults Across the United States: The STEP UP Randomized Clinical Trial. *JAMA Internal Medicine*, *179*(12), 1624–1632. https://doi.org/10.1001/jamainternmed.2019.3505

Popkin, M. L., & Ng, D. Y. (Eds.). (2022). *Future thinking in Roman culture: New approaches to history, memory, and cognition*. Routledge.

Raaijmakers, S. F., Baars, M., Schaap, L., Paas, F., & van Gog, T. (2017). Effects of performance feedback valence on perceptions of invested mental effort. *Learning and Instruction*, *51*, 36–46. https://doi.org/10.1016/j.learninstruc.2016.12.002

Racsmány, M., Demeter, G., & Szőllősi, Á. (2019). Successful list-method directed forgetting without retroactive interference of post-instruction learning. *Memory*, *27*(2), 224–230. https://doi.org/10.1080/09658211.2018.1499936

Rangel, A., & Clithero, J. A. (2012). Value normalization in decision making: Theory and evidence. *Current Opinion in Neurobiology*, *22*(6), 970–981. https://doi.org/10.1016/j.conb.2012.07.011

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

Risko, E. F., Medimorec, S., Chisholm, J., & Kingstone, A. (2014). Rotating with rotated text: A natural behavior approach to investigating cognitive offloading. *Cognitive Science*, *38*(3), 537–564. https://doi.org/10.1111/cogs.12087

Rothen, N., & Meier, B. (2010). Grapheme-colour synaesthesia yields an ordinary rather than extraordinary memory advantage: Evidence from a group study. *Memory*, *18*(3), 258–264. https://doi.org/10.1080/09658210903527308

Rouault, M., Dayan, P., & Fleming, S. M. (2019). Forming global estimates of self-performance from local confidence. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-019-09075-3

Runge, Y., Frings, C., & Tempel, T. (2019). Saving-enhanced performance: Saving items after study boosts performance in subsequent cognitively demanding tasks. *Memory*, *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*, cid:a028af1a-6b5a-4c14-acbf-

516ac19bdd78. https://doi.org/10.1007/s00426-020-01341-0

Sachdeva, C., & Gilbert, S. J. (2020). Excessive use of reminders: Metacognition and effort-minimisation in cognitive offloading. *Consciousness and Cognition*, *85*, 103024. https://doi.org/10.1016/j.concog.2020.103024

Sahakyan, L. (2004). Destructive effects of “forget” instructions. *Psychonomic Bulletin and Review*, *11*(3), 555–559. https://doi.org/10.3758/BF03196610

Sahakyan, L., Delaney, P. F., Foster, N. L., & Abushanab, B. (2013). List-method directed forgetting in cognitive and clinical research: A theoretical and methodological review. In *Psychology of learning and motivation - Advances in research and theory* (Vol. 59, pp. 131–189). Academic Press Inc. https://doi.org/10.1016/B978-0-12-407187-2.00004-6

Saunders, B., Lin, H., Milyavskaya, M., & Inzlicht, M. (2017). The emotive nature of conflict monitoring in the medial prefrontal cortex. *International Journal of Psychophysiology*, *119*, 31–40. https://doi.org/10.1016/j.ijpsycho.2017.01.004

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human Computer Studies*, *45*(2), 185–213. https://doi.org/10.1006/ijhc.1996.0048

Scarampi, C., & Gilbert, S. J. (2020a). 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. (2020b). 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

Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers*. Houghton Mifflin.

Schillemans, V., Luwel, K., Bulté, I., Onghena, P., & Verschaffel, L. (2009). The influence of previous strategy use on individuals' subsequent strategy choice: Findings from a numerosity judgement task. *Psychologica Belgica*, *49*(4), 191–205. https://doi.org/10.5334/pb- 49-4-191

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/https://doi.org/10.1037/a0022842

Sharma, D., Booth, R., Brown, R., & Huguet, P. (2010). Exploring the temporal dynamics of social facilitation in the stroop task. *Psychonomic Bulletin and Review*, *17*(1), 52–58. https://doi.org/10.3758/PBR.17.1.52

Shekhar, M., & Rahnev, D. (2020). The nature of metacognitive inefficiency in perceptual decision making. *Psychological Review*, *128*(1), 45–70. https://doi.org/doi:10.1037/rev0000249

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*, *40*(1), 99–124. https://doi.org/10.1146/annurev-neuro-072116-031526

Smith, R. E., & Bayen, U. J. (2005). The effects of working memory resource availability on prospective memory: A formal modeling approach. *Experimental Psychology*, *52*(4), 243–256. https://doi.org/10.1027/1618-3169.52.4.243

Soares, J. S., & Storm, B. C. (2018). Forget in a flash: A further investigation of the photo-taking-impairment effect. *Journal of Applied Research in Memory and Cognition*, *7*(1), 154–160. https://doi.org/10.1016/j.jarmac.2017.10.004

Song, C., Kanai, R., Fleming, S. M., Weil, R. S., Schwarzkopf, D. S., & Rees, G. (2011). Relating inter-individual differences in metacognitive performance on different perceptual tasks. *Consciousness and Cognition*, *20*(4), 1787–1792. https://doi.org/10.1016/j.concog.2010.12.011

Sparrow, B., & Chatman, L. (2013). Social Cognition in the Internet Age: Same As It Ever Was? *Psychological Inquiry*, *24*(4), 273–292. https://doi.org/10.1080/1047840X.2013.827079

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, *333*(6043), 776–778. https://doi.org/10.1126/science.1207745

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

Storm, B. C., Stone, S. M., & Benjamin, A. S. (2017). Using the Internet to access information inflates future use of the Internet to access other information. *Memory*, *25*(6), 717–723. https://doi.org/10.1080/09658211.2016.1210171

Tufft, M. R. A., & Richardson, D. C. (2020). Social offloading: Just working together is enough to remove semantic interference. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*. Cognitive Science Society.

Tversky, A., & Simonson, I. (1993). *Context- dependent Preferences*. *39*(10), 1179–1189. https://doi.org/https://doi.org/10.1287/mnsc.39.10.1179

Tversky, B. (2011). Visualizing thought. *Topics in Cognitive Science*, *3*, 499–535. https://doi.org/10.1111/j.1756-8765.2010.01113.x

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive*

Psychology, *71*, 1–26. https://doi.org/j.cogpsych.2014.01.003

Vaghi, M. M., Luyckx, F., Sule, A., Fineberg, N. A., Robbins, T. W., & De Martino, B. (2017). Compulsivity Reveals a Novel Dissociation between Action and Confidence. *Neuron*, *96*(2), 348-354.e4. https://doi.org/10.1016/j.neuron.2017.09.006

Vortac, O. U., Edwards, M. B., & Manning, C. A. (1995). Functions of external cues in prospective memory. *Memory*, *3*(2), 201–219. https://doi.org/10.1080/09658219508258966

Vorwerk, T. M., Shelton, S. T., & Hood, R. W. (2022a). Keeping your glass half full: Cognitive strategy intervention for older adults. *Gerontology and Geriatric Medicine*, *8*, 1–8. https://doi.org/10.1177/23337214221089178

Vorwerk, T. M., Shelton, S. T., & Hood, R. W. (2022b). Keeping Your Glass Half Full: Cognitive Strategy Intervention for Older Adults. *Gerontology and Geriatric Medicine*, *8*, 233372142210891. https://doi.org/10.1177/23337214221089178

Walsh, M. M., & Anderson, J. R. (2009). The strategic nature of changing your mind. *Cognitive Psychology*, *58*(3), 416–440. https://doi.org/10.1016/j.cogpsych.2008.09.003

Wegner, D. M. (1995). A computer network model of human transactive memory. *Social Cognition*, *13*(3), 319–339. https://doi.org/10.1521/soco.1995.13.3.319

Weis, P. P., & Wiese, E. (2020). Know your cognitive environment! Mental models as crucial determinant of offloading preferences. *Human Factors*. https://doi.org/10.1177/0018720820956861

West, J. T., & Mulligan, N. W. (2019). Prospective metamemory, like retrospective metamemory, exhibits underconfidence with practice. *Journal of Experimental Psychology: Learning Memory and Cognition*, *Advance online publication*.

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), e68210. https://doi.org/10.1371/journal.pone.0068210

Wilson, B. A., Emslie, H. C., Quirk, K., & Evans, J. J. (2001). Reducing everyday memory and planning problems by means of a paging system: A randomised control crossover study. *Journal of Neurology Neurosurgery and Psychiatry*, *70*(4), 477–482. https://doi.org/10.1136/jnnp.70.4.477

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, *9*(4), 625–636. https://doi.org/10.3758/BF03196322

Yaro, C., & Ward, J. (2007). Searching for Shereshevskii: What is superior about the memory of synaesthetes? *Quarterly Journal of Experimental Psychology*, *60*(5), 681–695. https://doi.org/10.1080/17470210600785208

Yates, F. A. (1966). *The art of memory*. University of Chicago Press.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1310–1321. https://doi.org/10.1098/rstb.2011.0416