The effect of recent reminder setting on subsequent strategy and performance in a prospective memory task

Chiara Scarampi[a]* and Sam J. Gilbert[a]

[a]*Institute of Cognitive Neuroscience, University College London, United Kingdom*

Address correspondence to:
Chiara Scarampi
Institute of Cognitive Neuroscience
17 Queen Square
London WC1N 3AR
UK
Email: c.scarampi@ucl.ac.uk
Tel: +44 (0)20 7679 1199

The effect of recent reminder setting on subsequent strategy and performance in a prospective memory task

The technological advancement that is rapidly taking place in today's society allows increased opportunity for "cognitive offloading" by storing information in external devices rather than relying on internal memory. This opens the way to fundamental questions regarding the interplay between internal and external memory and the potential benefits and costs of placing information in the external environment. This article reports the results of three pre-registered online experiments investigating the consequences of prior cognitive offloading on A) subsequent unaided ability, and B) strategic decisions whether to engage in future cognitive offloading. We administered a web-based task requiring participants to remember delayed intentions for a brief period and manipulated the possibility of setting reminders to create an external cue. Earlier cognitive offloading had little effect upon individuals' subsequent unaided ability, leading to a small and nonsignificant drop in subsequent performance. However, there was a strong effect on participants' subsequent likelihood of setting reminders. These findings suggest that the short-term impact of cognitive offloading is more likely to be seen on individuals' strategy choices rather than basic memory processes.

Keywords: Prospective Memory; Cognitive Offloading; Technology; Reminders; Delayed Intentions; Strategy Use

Introduction

In everyday life, we often form intentions for future actions which can only be executed after a delay. The ability to remember to perform such actions in the future is known as *prospective memory* (Einstein & McDaniel, 1990; McDaniel & Einstein, 2007). However, our mental abilities have acknowledged limits (e.g., we can encode and store in memory a limited amount of information; Cowan, 2010), so that we sometimes forget to execute an intended behaviour. When establishing an intention to act later, we can alternatively choose to use physical actions (e.g., setting a reminder such as an alarm or calendar alert) as a means of "cognitive offloading" (Risko & Gilbert, 2016) so that the information processing requirements of the task are altered. This in turns reduces cognitive demand and relieves the burden of intention maintenance in internal memory (Herrmann et al., 1999).

Although the phenomenon is ancient, researchers have only recently started conducting systematic investigations on the topic. This has been fuelled by rapid technological advancement, which allows us to use the environment (i.e., external memory) to record information rather than storing it in our brain (i.e., internal memory; Finley et al., 2018). Researchers have started questioning what the effects of cognitive offloading are on our future thinking and behaviour. In the current study, we explore the effects of recent use of reminders on subsequent performance and strategy in a task requiring memory for delayed intentions. More precisely, we study the effect that reminder setting (also known as "intention offloading"; Gilbert, 2015a) has on unaided memory ability and whether it influences the propensity to use the same strategy for future performance.

The relationship between intention offloading and unaided memory

While some experiments have shown that offloading information can improve performance when the external store is available (see Risko & Gilbert, 2016), other research has documented other potential effects of relying on external devices on cognition. Fisher, Goddu, and Keil (2015) documented an illusion of knowledge by which individuals using the Internet to search for information tend to conflate information available online with information stored in their own memory. In a similar vein, Ward (2013) found that using the internet causes people to assimilate its attributes into the self, overlooking the effect of the internet on performance and inflating judgements about the ability to perform well in the future.

Other research has documented instances where relying on external memory may prevent individuals from maintaining an internal representation of the information they are attempting to preserve (Henkel, 2014; Soares & Storm, 2018; Tamir et al., 2018). Henkel (2014) documented a photo-taking-impairment effect by which participants may be less likely to remember objects they photograph rather than objects they simply observe. Photography can be interpreted as a form of cognitive offloading, which allows individuals to rely on the external device to record information about the object, rather than storing it in the internal memory. This result has been recently replicated by Soares and Storm (2018), who found that participants exhibit a photo-taking-impairment effect even when they do not expect to have subsequent access to the photos. Converging support has been provided by a further study showing that media usage can impair memory (Tamir et al., 2018). Altogether, these findings suggest that the act of cognitive offloading may have a detrimental effect on internal memory for the offloaded information.

A common concern is that relying on cognitive aids can hinder unaided ability (Baldwin et al., 2011) by reducing opportunities for individuals to develop and maintain the skills required to perform tasks in an unaided manner (i.e., 'use it or lose it'). According to this view, simplifying a task by adopting a cognitive offloading strategy would lead to impaired subsequent performance on a more demanding task (i.e., one which does not afford offloading).

An alternative view is offered by the literature on ego depletion (see Friese et al., 2018 for a review). Although the literature in the field is controversial, previous research has suggested that initially performing a demanding self-control task can lead to impaired subsequent cognitive performance, compared with initial performance of an easier task. While some authors propose that this is because self-control is a depletable resource (Baumeister et al., 1998), others have challenged this view both conceptually (Lurquin & Miyake, 2017) and empirically (Hagger et al., 2016). For the present purposes, we simply note that the ego depletion literature makes an opposite prediction to the "use it or lose it" account, namely that performing a simplified task by adopting a cognitive offloading strategy might improve subsequent cognitive performance in a demanding task rather than impairing it. This prediction does not imply acceptance of the concept of resource depletion and would be equally compatible with other models of ego depletion (e.g., the idea that these effects reflect participants' beliefs and motivational states, rather than self-control being a limited resource).

Previous work has investigated the effect that saving a file on a computer has on the remembering of new information (Storm & Stone, 2015). Not only is the proportion of information recalled from the saved file higher than the information recalled when the file is not saved, but also, saving a file before studying a new file improves recall of the contents of the new file. That is, saving files represents a means to offload to-be remembered information onto outside sources, which facilitates the later encoding and remembering of new information. This effect suggests that offloading to-be-remembered information can improve learning of subsequent information. This finding was replicated in a recent study by Runge Frings, and Tempel (2019) who found that the benefits of memory offloading are not limited to memory performance but can free cognitive resources that can be used for subsequent unrelated tasks. Of the two accounts described above, this would be more compatible with the ego depletion model. However, in the paradigm used by Storm and Stone and Runge and colleagues, participants must learn the second file whilst still needing to remember the first one, potentially leading to interference and competition between the two files when the first one is not offloaded.

By contrast, in this study we were interested in investigating whether initially adopting an offloading or a non-offloading strategy affects subsequent memory performance, even when the original material has already been tested and is no longer relevant. In other words, we tested whether cognitive offloading at time A affects cognitive performance at time B, even when the material learned at time A is no longer relevant in the subsequent phase of the task. We did this in Experiment 1 by adapting the intention offloading paradigm used by Gilbert (2015a). Participants performed a memory test in two phases. In the first phase participants were randomised into two groups performing the task either with external reminders or unaided memory. In the second phase, all participants performed using unaided memory. This allowed us to test the impact (if any) of using external reminders on subsequent unaided memory performance.

The relationship between previous reminder setting and subsequent strategy choice

Beside understanding of the consequences of cognitive offloading on our unaided memory ability, another fundamental issue is the investigation of the factors that affect the likelihood of choosing to offload cognition rather than rely on the internal memory. In recent years, researchers have focused on different factors that affect cognitive offloading, such as task difficulty (Gilbert, 2015a), metacognitive beliefs (Gilbert, 2015b), age (Gilbert, 2015a), task instructions (Boldt & Gilbert, 2019), feedback valence and practice-trial difficulty (Gilbert et al., 2019). A further potential factor that has not been studied until now is participants' past history and previous experience with the act of cognitive offloading.

A growing body of research has shown that when solving a problem, individuals are inclined to reuse the strategy they used before (Luchins, 1942; Schillemans et al., 2010). This effect has been demonstrated in several domains ranging from anagrams (Ellis & Reingold, 2014) to perceptual decision-making (Schillemans et al., 2010) and chess (Bilalić et al., 2008). A similar *perseveration effect* has been found when individuals use the Internet as a form of cognitive offloading (Ferguson et al., 2015; Storm et al., 2016). Ferguson and colleagues (2015) found that access to the Internet influences individuals' willingness to volunteer self-generated answers, leading to fewer correct answers overall but greater accuracy when an answer was offered. In a similar vein, Storm and colleagues (2016) found that using Google to answer an initial set of trivia questions make individuals more likely to rely on the Internet to answer a new, relatively easier, set of questions. This suggests that once individuals are in the habit of using the Internet, they keep using it, even when it is not needed.

Starting from these observations that behaviour is influenced by the strategies previously used, another aim of this study was to test the hypothesis that the use of reminders in an offloading task increases the likelihood of relying on this strategy in the future. We did this in Experiments 2 and 3 by adapting the intention offloading paradigm used by Gilbert (2015a). Participants performed a memory test in two phases. In the first phase participants were randomised into two groups performing the task either with external reminders or unaided memory. In the second phase, all participants were given the opportunity to choose whether to use reminders or rely on their own memory. This allowed us to test whether once people are in the habit of using reminders, they keep using them.

Experiment 1

This experiment aimed to investigate whether initially adopting an offloading or a non-offloading strategy affects subsequent unaided memory performance, even when the original material has already been tested and is no-longer-relevant.

We evaluated the evidence for the following three alternative hypotheses:

- The use of external reminders impairs subsequent unaided memory performance, as predicted by the 'use it or lose it' account.
- The use of reminders improves subsequent unaided ability to execute a future intention, as predicted by the 'ego depletion' account.
- The use of external reminders has no impact on subsequent unaided memory performance.

Before commencing data collection, we preregistered our hypotheses, sample size, experimental procedure, participant exclusion criteria, and analysis plans (https://osf.io/724x5/).

Method

Participants

A total of 220 participants were recruited from the Amazon Mechanical Turk website (http://www.mturk.com), an online marketplace in which participants receive payment for completion of web-based tasks (Crump et al., 2013).

A statistical power analysis was performed with G*Power 3.1 for sample size estimation. To the best of our knowledge, no previous studies have investigated the effect of reminder use on subsequent unaided memory. As a consequence, for the purpose of sample size estimation, we referred to the literature on ego depletion, as similarly to our experiment, the typical paradigm used to test ego depletion consists of two conditions, both requiring participants to complete two consecutive tasks. We based the computations on a recent meta-analysis of the ego depletion effect conducted by Dang (2018). Since none of the tasks used in the ego depletion literature is comparable to that used in our experiment (i.e., a delayed intention task coupled with an arithmetic interruption question), we referred to the overall effect found in the meta-analysis: g = 0.38. With an alpha = .05 and power = 0.80, the projected sample size needed with this effect size was approximately N = 220 for the simplest between group comparison (110 participants in each group).

As in earlier studies by Gilbert (2015a, 2015b), participation was restricted to volunteers aged at least 18 years and living in the USA, to reduce heterogeneity. Furthermore, a total of 23 participants did not meet the criteria of at least 80% accuracy in the arithmetic-verification test and/or 50% accuracy in the intention-offloading task and were therefore excluded and replaced.

Participants were randomly allocated to two groups: Forced Reminders and No Reminders. In the *No Reminders* condition there were 67 male and 43 female participants with a mean age of 36±10 years (range 22-63). In the *Forced Reminders* condition there were 62 male and 48 female participants with a mean age of 36±10 years (range 19-70). Participation took on average 25 minutes (minimum = 8 minutes, maximum = 2 hours 3 minutes) in the No Reminders condition and 25 minutes (minimum = 10 minutes, maximum = 3 hours 12 minutes) in the Forced Reminders condition,[1] for which participants were paid $2.50. Ethical approval was received from the local ethics committee and participants provided informed consent before participating in the study.

[1] The completion time refers to the duration between the participant first opening the task in their web-browser and their final completion. Participants were not necessarily working on the task for this full duration seeing as they were free to take breaks at several points. See Supplementary Materials for analyses investigating the relationship between completion time and task performance.

Task and Procedure

The 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 the Eclipse development environment (https://www.eclipse.org).

Intention-offloading task. In each trial, ten yellow circles numbered 1 to 10 were randomly positioned within a square. Participants were instructed to use the mouse to drag the circles in ascending order (1, 2, 3, etc.) to the bottom of the square. Each time a square was dragged to the bottom, it disappeared, leaving the other circles on the screen. After the 10th circle disappeared, the screen was cleared and the next trial began (see Figure 1 for a schematic illustration of the task and visit http://www.ucl.ac.uk/sam-gilbert/demos/circleDemo.html for a demonstration).

Alongside this ongoing task, on each trial participants were provided with delayed intentions. That is, they were instructed to drag three circles to a specific alternative location (i.e., left, right, or top of the square) when the numbers were reached in the sequence.[2] This led to the formation of delayed intentions to perform particular actions when they encountered prespecified cues, although participants could produce a standard ongoing response (i.e., dragging the circle to the bottom of the box) if they forgot.

The task also permits intention offloading in a simple manner: at the beginning of each trial, participants can drag the target circles towards their intended location.

[2] In each trial, participants had to drag one target circle to the left, one target circle to the top, and one target circle to the right. The three target numbers were randomly chosen from the range 3-10 and randomly associated with the three target sides. When the instructions were presented at the beginning of each trial, the instructions for the three target circles were presented in ascending order.

From this point on, there is no need to mentally rehearse the delayed intention(s). Instead, the locations of the target circles themselves represent the intention, providing a perceptual trigger when they are reached in the sequence. An everyday analogue might be leaving an object by the front door, so that we remember to take it with us when leaving the house. [Figure 1 approximately here]

There are three potential intention offloading conditions in this paradigm. In the No Reminders condition, participants rely on internal memory only. This is enforced by fixing the positions of all circles on screen apart from the next in the sequence, so that it is not possible to adjust the position of forthcoming targets when they first appear on screen. In the Forced Reminders condition, participants are required to set external reminders. This is enforced by preventing participants from continuing the ongoing task (removing circles from the bottom of the box) until they have adjusted the position of any new target circles that appear. In the Optional Reminders condition (used in Experiments 2 and 3), participants have a free choice whether or not to set reminders.

Arithmetic-verification test. Participants in both conditions additionally received a distracting arithmetic question during each trial, via a pop-up box (using the same procedure as Gilbert, 2015a). This occurred immediately after dragging one of the nontarget circles to the bottom of the box, at a position in the sequence randomly selected between the first circle and the circle immediately before the first target. This ensured that participants always had the opportunity to set reminders before the first arithmetic question was presented. In line with the literature showing that a brief task interruption can tax cognitive resources and lead to errors (e.g., Weakley & Schmitter-Edgecombe, 2019), we introduced the arithmetic-verification test to increase the difficulty of the task, helping to reduce the possibility of ceiling effects in the measure of intention fulfilment.

Manipulation. The experiment consisted of 20 experimental trials, divided in two phases and the key element of the study was the manipulation of intention offloading in the first phase (See Figure 2 for a schematic representation of the experimental design). [Figure 2 approximately here] More precisely, at the beginning of the experiment, participants were randomly assigned to one of two conditions: No Reminders or Forced Reminders. Participants in the No Reminders condition were asked to complete the intention offloading task relying on their internal memory only, whereas participants in the Forced Reminders condition were asked to complete the task making use of reminders.

A second phase of the experiment followed immediately after the first and was identical for participants in both conditions. That is, they were given 10 trials of the same intention-offloading task, but this time none of them was allowed to set reminders. For the exact wording of all experimental instructions, see Supplementary Materials. Furthermore, the full source code for running the experiment can be downloaded at https://osf.io/3f52c/.

Apparatus

Participants completed the task via their computer's web browser. As in Gilbert (2015a), participation was only permitted if the browser window had dimensions of at least 500 x 500 pixels. The square box containing the circles was sized at 80% of the horizontal or vertical extent of the browser window, whichever was smaller. Each circle had a radius of 5.5% of the width/height of the box, and all circles were initially placed so that they fell within a central portion of the box with dimensions sized at 56% of the total width/height, so that no circles were adjacent to any of the edges of the box at the beginning of the trial.

Data Analysis

All data analyses were conducted in R version 3.5.2. Bayesian independent samples t-tests were conducted in JASP version 0.9.2. The key dependent variable was *target accuracy* – i.e., the proportion of targets that are correctly dragged to the instructed location rather than to the bottom of the square.

Results

Mean arithmetic-verification accuracy was 99%. We first compared accuracy between the two groups in phase 1, using a Welsh two-sample t-test. Accuracy was significantly higher in the Forced Reminders group than the No Reminders group, $t(210.07) = 3.69$, $p < .001$, $d = 0.50$ showing that the use of external reminders improves task performance compared to the use of internal memory only (see Figure 3). [Figure 3 approximately here]

Next, we investigated the effect of using reminders in the first phase on unaided memory ability in the second phase. This was our main hypothesis-testing analysis. An independent sample t-test revealed no reliable difference between the two groups, $t(218) = 1.03$, $p = .30$, $d = 0.14$.[3]

We complemented the classical null hypothesis significance testing (NHST) reported above with Bayesian analyses, which can provide a more informative and effective approach for hypothesis testing (Kruschke, 2013). More precisely, we used Bayes factors, which allowed us to quantify relative evidence for both alternative and null hypotheses (Quintana & Williams, 2018). These Bayes factors can be interpreted as a measure of the strength of evidence in favour of one theory among two competing

[3] A further exploratory analysis was performed to check for learning effects over time in the No Reminders group, however there was little evidence for this (see Supplementary Materials).

theories. These values are commonly interpreted using approximate guidelines such as Jeffery's scheme, which categorises Bayes Factors as anecdotal, moderate, strong, very strong, or extreme evidence for a hypothesis over another (Jeffreys, 1961).

We conducted three separate Bayesian t-test tests, corresponding to each of the three hypotheses above. A first test evaluating the evidence in favour of the hypothesis that the use of external reminders impairs subsequent unaided prospective memory performance ('use it or lose it' account) yielded a *BF* of 2.47 (anecdotal evidence) in favour of the null hypothesis that there is no effect of cognitive offloading on subsequent unaided memory ability. A second test evaluating the evidence in favour of the hypothesis that the use of reminders improves subsequent unaided ability to execute a future intention ('ego depletion' account) yielded a *BF* of 12.92 (strong evidence) in favour of the null hypothesis of no effect of reminder setting on unaided ability. Finally, a third test evaluating the evidence in favour of the hypothesis that use of external reminders has no impact on subsequent unaided memory performance yielded a *BF* of 4.14 (moderate evidence) in favour of the hypothesis that the No Reminders group and the Forced Reminders group can be best described by the same distribution.

Finally, we compared accuracy in each group between phase 1 and phase 2, using a paired t-test. Results showed a significant decrease from phase 1 to phase 2 in the Forced Reminders group ($t(109)=4.24$, $p < .001$, $d = 0.40$) and no significant difference in the No Reminders group ($t(109)=1.22$, $p = .223$, $d = 0.12$).[4]

[4] All the analyses were replicated including age as a covariate (see Supplementary Materials). There was little evidence for an effect of age on our results. The inclusion of age in the analyses also had little impact on our main hypothesis-testing analyses, and results were generally consistent between the two approaches.

Discussion

The results replicated the finding that individuals who set reminders fulfil their delayed intentions more often (Gilbert, 2015a). Furthermore, unaided memory ability in phase 2 did not significantly differ between the No Reminder and Forced Reminder groups. Therefore, prior use of a cognitive offloading strategy did not have a significant effect on subsequent unaided memory ability. There was strong evidence against the hypothesis that use of reminders improves subsequent unaided ability to execute a future intention (‘ego depletion’ account), and anecdotal evidence against the hypothesis that use of reminders harms subsequent unaided ability (‘use it or lose it’ account). Therefore, our results seem compatible with the idea that prior reminder setting has no impact on subsequent unaided performance, or a small detrimental effect (e.g., because it reduced practice opportunities; cf Baldwin et al., 2011). However, the possibility of a detrimental effect rests on a nonsignificant finding, so it is not strongly supported. Of course, our results can only speak to the impact of offloading on a short timescale, seeing as the total experiment duration was only about 25 minutes. It is quite possible that impact over a longer timescale could be different.

Our results contrast with previous evidence showing that prior offloading can *improve* subsequent unaided memory (e.g., Storm & Stone, 2015). We suggest that the key difference is that in Storm and Stone’s study, participants needed to maintain the original information (either in internal memory or an external store) while the subsequent memory test took place. Thus, participants had to learn new information whilst simultaneously maintaining earlier information either in internal memory or an external store. By contrast, in our study the earlier memoranda were no longer relevant at the time of the subsequent memory test. Therefore, regardless of whether the original information was stored internally or externally, it could be dismissed from memory by the time the subsequent memory test took place.

In summary, our results suggest that short-term use of a cognitive offloading strategy has little effect on memory ability in circumstances where the previously memorised information is no longer required and participants are forced to use unaided memory. We now turn to the question of whether prior cognitive offloading influences subsequent behaviour when participants are given a choice whether or not to set reminders.

Experiment 2

In Experiment 1, we investigated whether initially adopting an offloading versus a non-offloading strategy affects subsequent memory performance when participants are prevented from setting reminders. However, in most everyday life situations, individuals are free to choose whether or not to use a cognitive offloading strategy to reduce the cognitive demands of a task (Risko & Gilbert, 2016). This experiment aimed at understanding whether the use of a cognitive offloading strategy becomes habitual. We did this by adapting the intention offloading paradigm used in the previous experiment. Again, participants performed a memory test in two phases. In the first phase they were randomised into groups performing the task either with external reminders or unaided memory. In the second phase, all participants were given a free choice whether to use reminders or rely on their own memory.

We predicted that participants who were instructed to set reminders in the first phase of the experiment would rely more on reminders in the second phase than would participants who were instructed to solve the task using their own memory in phase 1. We also expected to replicate the finding that the use of reminders increases performance accuracy. Accordingly, we predicted that in phase 1 participants who were initially instructed to set reminders would be more accurate than participants who were initially asked to solve the task using their own memory.

Given that reminder-use is associated with accuracy, it might be predicted that the groups would differ in accuracy as well as reminder-use in phase 2. However, whereas phase 1 had a stark difference in reminder use between the groups (100% versus 0%), any difference in phase 2 was expected to be smaller. Our initial calculations suggested that sufficient power to detect any resulting influence on accuracy would require a large sample size (N > 3000). Therefore, although we report analyses of phase-2 accuracy, we did not expect our experiment to have sufficient power to detect any potential group difference on this measure, and our main focus is on the reminder-setting measure.

Before commencing data collection, we preregistered our hypotheses, sample size, experimental procedure, participant exclusion criteria and analysis plans (https://osf.io/9wtd8/).

Method

Participants

A total of 192 participants took part in the experiment. This was based on a power calculation aiming for 80% power to detect a between-group difference with effect size d = .41, based on unpublished pilot data (see https://osf.io/9wtd8/ for further information). This was our predicted effect size for the difference in reminder setting between groups who had previously performed the Forced Reminders versus the No Reminders versions of the task.

Participants were randomly allocated to two groups: Forced Reminders (96 participants) and No Reminders (96 participants). In the *Forced Reminders* condition there were 60 male and 36 female participants with a mean age of 35±11 years (range 19-68). In the *No Reminders* condition there were 62 male and 34 female participants with a mean age of 35±10 years (range 20-73). Participation took on average 21 minutes (minimum = 8 minutes, maximum = 2 hours 15 minutes) in the No Reminders condition and 23 minutes (minimum = 10 minutes, maximum = 2 hours 5 minutes) in the Forced Reminders condition, for which participants were paid $2.50. The same exclusion criteria as in experiment 1 were used. A total of 13 participants were excluded and replaced because they did not meet the inclusion criteria. All participants provided informed consent before participating and the research was approved by the UCL Research Ethics Committee.

Data Analysis

The analyses focused on two key dependent measures: *target accuracy* – i.e., the proportion of targets that are correctly dragged to the instructed location rather than to the bottom of the square – and *externalising proportion* – i.e., the proportion of target circles for which participants set up an external reminder, by moving it to a different location before reaching its position in the ongoing task. For each trial, a record was made of any time that a participant clicked on a circle when it was not next in sequence (e.g., clicking on circle number 5, when circle number 1 is still on the screen). The proportion of targets clicked before their turn in the sequence was then calculated, by dividing the number of target circles that were clicked when it was not their turn in the sequence by the total number of targets (i.e., three). An analogous procedure was used to calculate the proportion of nontarget circles that were clicked out of sequence. This was subtracted from the target-circle proportion, to control for any out-of-sequence clicks that might have occurred simply due to accidentally clicking on the wrong circle, regardless of its status as target or nontarget. This yielded the externalizing proportion, calculated in the same manner as earlier studies (e.g., Gilbert, 2015a) so that 0 would indicate that participants did not click on target circles out of sequence any more often than nontarget circles, and where 1 would indicate that participants clicked every target circle before reaching its position in the sequence, but never clicked on nontarget circles in this way.

Results

Mean arithmetic-verification accuracy was 99%. Participants rarely clicked on nontarget circles when it was not their turn in the sequence (< 5%). Therefore, the externalising proportions described below predominantly reflect participants' tendency to click target circles before it was their turn in the sequence.

Objective Accuracy

We first compared accuracy between the two groups in phase 1, using a Welsh two-sample t-test. Accuracy was significantly higher in the Forced Reminders group than the No Reminders group, $t(180.39) = 4.39$, $p < .001$, $d = 0.63$ showing, as in the previous experiment, that the use of external reminders improves task performance compared with the use of internal memory only (see Figure 3).

Next, we conducted an independent samples t-test comparing accuracy in phase 2 between the two groups. We did not find any significant difference between the two groups, $t(190)=0.06$, $p = .95$, $d = 0.01$. Finally, we compared accuracy in each group between phase 1 and phase 2, using a paired t-test. Results showed a marginally significant decrease from phase 1 to phase 2 in the Forced Reminders group ($t(95)=1.95$, $p = .053$, $d = 0.20$) and highly significant increase in the No Reminders group ($t(95)=3.34$, $p = .001$, $d = 0.34$).

Reminder Setting

As our main hypothesis-testing analysis, we investigated the effect of using reminders in the first phase on the likelihood of using reminders in phase 2. A Welsh two-sample t-test revealed a significantly higher externalisation proportion in the Forced Reminders group, $t(183.29) = 4.48$, $p < .001$, $d = 0.65$ (see Figure 4). [Figure 4 approximately here] A Bayesian independent samples t-test was also conducted to investigate the strength of the evidence in favour of our hypothesis against the null hypothesis. The test yielded a Bayes factor of 1315 (extreme evidence) in favour of the hypothesis that the use of reminders affects the likelihood of using them again in the future.

Relationship between Accuracy and Reminder Setting

Next, we ran a correlation between accuracy in phase 1 and reminder choice in phase 2 to investigate whether there is a relationship between performance in phase 1 and the choice to set reminders in phase 2. We found a significant positive association between the two variables for the Forced Reminder group, $r_s(94) = .28$, $p = .006$, and a significant negative correlation for the No Reminders group, $r_s(94) = -.25$, $p = .016$.

We also combined over the two groups and ran a multiple regression where group was a factor. The results of the regression indicated the two predictors explained 10% of the variance ($R^2 = .10$, $F(2,189) = 10.21$, $p < .001$). It was found that group significantly predicted externalising proportion in phase 2 ($\beta = .28$, $p < .001$), whereas accuracy in phase 1 was not a significant predictor.

Last, we were interested in studying whether the use of reminders improves accuracy in phase 2. We performed a multiple regression with phase-2 accuracy as dependent variable and phase-1 accuracy and externalising proportion in phase 2 as independent variables. We conducted this analysis separately for the two groups. For the No Reminders group, the two predictors explained a significant amount of the variance in phase-2 accuracy ($R^2 = .30$, $F(2,93) = 20.04$, $p < .001$). Phase-1 accuracy significantly predicted accuracy in phase 2 ($\beta = .44$, $p < .001$), as did externalising proportion (β = 10.83, p < .001). Also for the Forced Reminders group the two predictors explained a significant amount of the variance (R^2 = .26, F(2,93) = 16.29, p < .001). Again, phase-1 accuracy significantly predicted accuracy in phase 2 (β = .40, p < .001), as did externalising proportion (β = 12.85, p < .001).

All analyses above were included in our original pre-registered analysis plan. In an additional post-hoc analysis which was not pre-registered, we explored how reminder setting varied across trials in the two groups (see Figure 5). [Figure 5 approximately here] From visual inspection, it appears that in both groups the use of reminders remained relatively constant over time, being significantly higher for the Forced Reminder group in every trial (p < .05). This result was confirmed by a further analysis. We fitted individual regression lines predicting the externalising proportion from trial number and compared the average slopes in the two groups. The difference between the two slopes was not statistically significant (see Figure 6; t(188) = 0.14, p = .89, d = 0.02). [Figure 6 approximately here]

Discussion

As in Experiment 1, the use of reminders in phase 1 led to a higher accuracy, confirming the benefit of cognitive aids on memory accuracy.

Another relevant result is the significant correlations between phase-1 accuracy and externalising proportions. For participants in the No Reminders group, objective performance in phase 1 was negatively related to intention offloading in phase 2, suggesting that the experience of making errors when solving the task using their own unaided memory – reflected in a lower phase-1 performance – may have triggered an increase in phase-2 intention offloading. Conversely, for the Forced Reminder group, objective performance in phase 1 was positively related to intention offloading in phase 2, suggesting that participants who experienced successful performance using an intention offloading strategy were more likely to continue with the same strategy in phase 2.

As foreseen from the power calculation, we did not detect an effect of setting reminders on accuracy in phase 2. Nevertheless, although nonsignificant, accuracy in phase 2 was slightly higher for the Forced Reminder group than the No Reminder group. It is possible that an alternative task, manipulating task difficulty, would show a stronger relationship between reminder-use and accuracy.

The results suggest that once participants find a memory strategy that works adequately, they are likely to keep using it. We found that participants who used reminders in the first half of the experiment were significantly more likely to rely on this strategy again when given a free choice, suggesting that once people are in the habit of using reminders, they keep using them. Experiment 3 seeks to further look into the mechanisms that might explain this result.

Experiment 3

In the previous experiment, we found that the use of a cognitive offloading strategy increased the likelihood of relying on such strategy in the future. However, two different factors might explain the results: a) preference to use the more practiced, initially-learned strategy, or b) simple repetition of the last strategy adopted, regardless of how well-practised it was in the initial learning phase. The main aim of this study was to distinguish between these two possibilities. In order to obtain an experimental paradigm able to maximise the difference between these two explanations, we repeated the procedure of Experiment 2 with one modification: after phase 1 participants performed a single trial with the opposite strategy before commencing phase 2 where they had free choice over which strategy to use. This means that just before the free-choice trials, participants who learned the task without reminders had used reminders on the previous trial, and vice-versa.

Different results were expected depending on whether participants were biased towards the strategy they used when first learning and practising the task, versus the strategy used on the immediately preceding trial. We evaluated evidence for two main possibilities:

- If participants were biased to repeat the strategy they used when they first learned the task, the group that used reminders in phase 1 would be more likely to use reminders when given a free choice in phase 2, despite the reversal-trial between the two phases where the strategies are reversed.
- If participants were simply biased to repeat the strategy used on the previous trial, the opposite effect would be expected, and participants would be more likely to repeat the strategy used on the reversal-trial.

We also considered the possibility of a combination of these effects, i.e., participants could be influenced both by the initial strategy and the reversal-trial strategy. These effects would be expected to yield opposite influences and could potentially cancel each other out. To evaluate evidence for this possibility we compared results from this experiment to our earlier study. We acknowledge that between-experiment comparisons are not as well-controlled as within-experiment comparisons, nevertheless we expected that the between-experiment analyses could provide additional useful information.

The use of reminders may affect not only the likelihood of using them again in the future, but also performance accuracy (as well as participants' beliefs about performance accuracy; see Boldt & Gilbert, 2019; Gilbert, 2015b; Hu et al., 2019). We expected to replicate the finding that the use of reminders increases performance accuracy. Accordingly, we hypothesised that in phase 1 participants who are initially instructed to set reminders will be more accurate than participants who are initially asked to solve the task using their own memory. As in Experiment 2, our main outcome measure in phase 2 was the externalising proportion, seeing as a larger sample size would be required to detect any group difference in accuracy. Before commencing data collection, we preregistered our hypotheses, sample size, experimental procedure, participant exclusion criteria, and analysis plans (https://osf.io/vg6tz/).

Method

Participants

A total of 192 participants took part in the experiment.

A statistical power analysis was performed with G*Power 3.1 for sample size estimation. The computations were based on the results of Experiment 2. This revealed an effect size of $d = 0.65$ when comparing the proportion of reminder setting between participants who in a first phase always used reminders with those who never did. With an alpha = .05 and power = 0.80, the projected sample size needed with this effect size is approximately $N = 76$ (38 participants in each group). However, we expected to find a smaller effect than in the previous study seeing as the reversal-trial manipulation may weaken or even reverse any effect, which would require a larger sample. Mindful of the intention to compare the results with those of the previous study, we decided to keep the same sample size: $N = 192$ (96 participants in each group).

Participants were randomly allocated to two groups: Forced Reminders (96 participants) and No Reminders (96 participants). In the *Forced Reminders* condition there were 55 male and 41 female participants with a mean age of 35±11 years (range 19-70). In the *No Reminders* condition there were 64 male and 32 female participants with a mean age of 35±10 years (range 21-68). Participation took on average 26 minutes (minimum = 11 minutes, maximum = 2 hours 15 minutes) in the No Reminders condition and 29 minutes (minimum = 11 minutes, maximum = 1 hour 44 minutes) in the Forced Reminders condition, for which participants were paid $2.50. The same exclusion criteria as in the previous experiments were used. A total of 7 participants were excluded and replaced because they did not meet these criteria. All participants provided informed consent before participating and the research was approved by the UCL Research Ethics Committee.

Results

Mean arithmetic-verification accuracy was 99%. As in Experiment 2, participants rarely clicked on nontarget circles when it was not their turn in the sequence (< 5%). Therefore, the externalising proportions described below predominantly reflect participants' tendency to click target circles before it was their turn in the sequence.

Objective Accuracy

We first compared accuracy between the two groups in phase 1, using a Welsh two-sample t-test. Accuracy was significantly higher in the Forced Reminders group than the No Reminders group, $t(161.7) = 3.89$, $p < .001$, $d = 0.56$ showing, as in the previous studies, that the use of external reminders improves task performance compared with the use of internal memory only (see Figure 3).

Next, we conducted an independent samples t-test comparing accuracy in phase 2 between the two groups. We did not find any significant difference between the two groups, $t(190)=0.21$, $p = .83$, $d = 0.03$.

Finally, we compared accuracy in each group between phase 1 and phase 2, using a paired t-test. There was a nonsignificant drop in accuracy from phase 1 to phase

2 in the Forced Reminders group ($t(95)=1.02$, $p = .31$, $d = 0.10$) and highly significant increase in the No Reminders group ($t(95)=4.71$, $p < .001$, $d = 0.48$).

Reminder Setting

As our main hypothesis-testing analysis, we first compared the likelihood of using reminders in phase 2 between the two conditions. An independent samples t-test found no significant differences in the externalising proportion in the two groups, $t(190) = 0.04$, $p = .97$, $d = 0.01$ (see Figure 4). A Bayesian independent samples t-test was also conducted to investigate the strength of the evidence in favour of our hypothesis against the null hypothesis. The test yielded a Bayes factor of 6.37 (moderate evidence) in favour of the null hypothesis postulating that there is no difference in reminder setting between the two conditions.

As mentioned above, we also considered the possibility that the strategy used in phase 1, as well as the strategy used on the reversal-trial, might both have influenced strategy choice in phase 2. These effects would operate in opposite directions and could potentially cancel each other out leading to a reduced effect of phase-1 strategy on phase-2 choices, or no effect at all. To explore evidence for this possibility we compared results from this experiment to those of Experiment 2 performing a 2 x 2 ANOVA with Condition and Experiment as factors. There was a significant effect of Condition, $F(1, 380) = 9.57$, $p = .002$, $\eta_p^2 = .025$, reflecting that on average, the externalising proportion was higher in the Forced Reminders condition than in the No Reminders group. There was also a significant interaction between Condition and Experiment, $F(1, 380) = 9.18$, $p = .003$, $\eta_p^2 = .024$, showing that the influence of initial phase-1 strategy on phase-2 reminder setting was significantly different between the experiments. The main effect of Experiment was not significant $F(1, 380) = 0.03$, $p = .85$, $\eta_p^2 = .00$.

Last, we performed an exploratory analysis to compare the externalising proportion in phase 2 across experiments separately for the two groups. For the No Reminders group, we found that the externalising proportion was marginally higher in Experiment 3 ($M = 0.60$) than in Experiment 2 ($M = 0.47$), $t(190) = 1.93$, $p = .056$, $d = 0.28$. The corresponding Bayesian independent samples t-test yielded a *BF* of 1.14 (anecdotal evidence) in favour of the null hypothesis of no difference between experiments. For the Forced Reminders group, the externalising proportion was significantly higher in Experiment 2 ($M = 0.74$) than in Experiment 3 ($M = 0.60$), $t(184.85) = 2.39$, $p = .018$, $d = 0.34$. The corresponding Bayesian independent samples t-test yielded a *BF* of 2.19 (anecdotal evidence) in favour of this cross-experiment difference.

Relationship between Accuracy and Reminder Setting

Next, we ran a correlation between accuracy in phase 1 and reminder choice in phase 2 to investigate whether there is a relationship between performance in phase 1 and choice to set reminders in phase 2. Correlation coefficients were positive in both groups but nonsignificant (No Reminders: $r_s(94) = .08$, $p = .44$; Forced Reminders: $r_s(94) = .12$, $p = .24$). We also combined over the two groups and ran a multiple regression where group was a factor. The results of the regression indicated the two predictors explained only 1% of the variance ($R^2 = .01$, $F(2,189) = 1.43$, $p = .24$).

Last, we were interested in studying whether the use of reminders improves accuracy in phase 2. We performed a multiple regression with phase-2 accuracy as dependent variable and phase-1 accuracy and externalising proportion in phase 2 as independent variables. We conducted this analysis separately for the two groups. For the No Reminders group, the two predictors explained a significant amount of the variance in phase-2 accuracy ($R^2 = .46$, $F(2,93) =39.64$, $p < .001$). Phase-1 accuracy significantly predicted accuracy in phase 2 (β = .46, p < .001), as did externalising proportion (β = 10.25, p < .001). Also for the Forced Reminders group the two predictors explained a significant amount of the variance (R^2 = .41, F(2,93) = 32.42, p < .001). Again, phase-1 accuracy significantly predicted accuracy in phase 2 (β = .70, p < .001), whereas externalising proportion was not a significant predictor (β = 1.57, p = .54).

Along with the pre-registered analyses above, we also conducted an exploratory analysis to examine how reminder setting varied across trials in the two groups (see Figure 5). The graph suggests that participants were initially more likely to rely on the strategy used in the reversal-trial and then returned to the strategy learnt originally in the first phase of the experiment. This result was confirmed by a further analysis. We fitted individual regression lines predicting the externalising proportion from trial number and compared the average slopes in the two groups. The difference between the two slopes was statistically significant (see Figure 6; t(188) = 2.51, p = .013, d = 0.36. Therefore, even though the groups did not differ significantly in the overall amount of reminder setting in phase 2, they did differ in the temporal profile of strategy choices over the 10 trials.

Discussion

Whereas phase-1 strategy had a strong effect on phase-2 reminder setting in Experiment 2, this effect seemed to be eliminated in the present experiment. This suggests that the strategy perseveration effect seen in the earlier experiment can be washed out by performing just one trial of the opposite strategy, or that the influences of phase-1 strategy (in one direction) and reversal-trial strategy (in the other) cancelled each other out. Either way, results indicate that the influence of the strategy used when first learning the task (seen in Experiment 2) can be modulated by the strategy used on a subsequent trial. Furthermore, the single-trial analysis showed that the groups differed in the dynamics of their strategy choices over time. This contrasts with the relatively stable strategy choices seen in Experiment 2. One interpretation of this would be that participants' strategy choices are influenced both by the immediately-preceding trial, and their earlier use of strategies when first learning the task. To test such an account, it would be necessary to manipulate the factors of phase-1 strategy and reversal-trial strategy independently, rather than manipulating them together as in the present experiment.

General Discussion

External memory is crucial in today's life, but it is also changing how people use their own memory systems (Nestojko et al., 2013). Two alternative arguments can be found in the literature. On the one hand, offloading memory onto the environment can overcome the limited capacity of working memory, relieve the burden on prospective memory, and improve cognitive functioning (Finley et al., 2018). On the other hand, externally stored information may not always be a viable replacement for internal memory and excessive reliance on external memory sources may reduce flexible access to internally stored knowledge, impairing in turn higher order cognition (Nestojko et al., 2013). Similar concerns date back at least to the time of Socrates, who worried over 2000 years ago that the use of writing would "introduce forgetfulness into the soul of those who learn it: they will not practice using their memory because they will put their trust in writing" (Plato, approximately 370 BC/1995, p. 79). To date, despite widespread interest in these issues (e.g., "Is Google making us stupid?"; Carr, 2008), relatively little empirical data has been collected to address the two possibilities outlined above. We therefore conducted the present study to collect further evidence in the context of intention offloading to better understand the effect that the use of external reminders has on unaided memory ability and subsequent strategy choice.

The first aim of this study was to investigate the influence of earlier offloading on subsequent unaided memory ability. Some previous research has suggested that the act of cognitive offloading may have a detrimental effect on unaided memory for the specific memories that are offloaded (Henkel, 2014; Soares & Storm, 2018; Tamir et al., 2018). Other studies suggest that offloading memory to an external store can improve individuals' ability to encode and remember subsequently-presented information (Storm & Stone, 2015). Here, we found no significant impact of earlier offloading on subsequent performance. We suggest that the key difference between our study and the earlier ones is as follows. In studies that report a detrimental effect of offloading on memory (Henkel, 2014; Soares & Storm, 2018; Tamir et al., 2018), participants were tested on their unaided recall of the very information that was either offloaded or maintained internally. In studies that report an improvement to memory caused by offloading (Storm & Stone, 2015), participants are tested on their unaided memory for new information acquired whilst still maintaining earlier information that was either offloaded or not. In both cases, the information that was originally offloaded or maintained internally continued to be task-relevant at the time of the main hypothesis-testing memory test. By contrast, our experiment compared two conditions where the earlier information was no longer required when the hypothesis-testing memory test took place, regardless of whether it was originally offloaded or maintained internally. Our results suggest that in these circumstances, short-term use of an offloading or internal strategy has little effect on subsequent unaided memory performance. Of course, it is quite possible that longer-term or habitual use of an offloading or internal memory strategy over an extended period of time might impact on unaided memory abilities. Investigating this question would require a different research design to the one used here.

Furthermore, although accuracy was relatively high in our experiments (leading to potential concerns about ceiling effects) all three experiments showed a significant effect of cognitive offloading on performance when it was manipulated as an experimental factor, replicating previous related findings (e.g., Gilbert, 2015b; Storm & Stone, 2015). In addition, Experiments 2 and 3 showed that individual differences in offloading strategies were associated with accuracy in phase 2. The finding that offloading strategies are associated with improved task performance is highly relevant when considering the interplay between internal and external memory. While much of the ongoing debate since the time of Socrates has focused on whether or not offloading has a detrimental effect on our unaided ability, it should also be recognised that *not* engaging in cognitive offloading could also have a detrimental effect on our performance, in that it can impair subsequent strategy choice, in contexts where an offloading strategy is available and effective.

The second aim of the current study was to address this issue, by testing whether individuals' strategy choice is affected by previous strategies used in the context of an intention offloading task. The results of Experiment 2 showed that once participants were in the habit of using reminders, they continued to rely on them when given a free choice. The findings align with previous research documenting the so-called *Einstellung effect* (Luchins, 1942) or perseveration effect (Schillemans et al., 2010) showing that the repeated implementation of a strategy increases the likelihood of using the same strategy again. This effect might explain why in the second phase of the experiment participants in both conditions were biased towards the strategy used in the earlier phase. That is, this mechanism made participants blind to the possibility of using another strategy, even a more efficient one, such as switching to the use of reminders for the No Reminders group. Support for this interpretation comes from the literature on the priming effect – that is, the recent use of a specific cognitive strategy leads to a temporary increase in the likelihood of applying the same procedure in a subsequent task (Higgins, 1996). As a consequence, when given a free choice in phase 2, participants might have favoured the primed strategies – i.e., the one used in phase 2 - at the expense of the other one.

Furthermore, there is evidence showing that after switching strategy in a cognitive task, participants exhibit longer reaction times and possibly higher error rates than when repeating the same strategy (Lemaire & Lecacheur, 2010; Luwel et al., 2009). It is possible that in the second phase of experiment 2 participants continued to apply the strategy learnt in phase 1 in the attempt to avoid such a strategy-switch cost. As noted by Schillemans and colleagues (2012), in certain situations, such as when different strategies are almost equally valid, continuing to apply the same strategy might be more adaptive than switching to another one because the strategy-switch cost may be higher than the benefit obtained by the use of the more efficient strategy. This might explain why participants in the No Reminders group kept solving the task using their own memory, even though switching to the use of reminders could have been more effective. Further research might be conducted manipulating the difficulty level of the task to study the threshold between strategy-switch cost and reminder-setting benefits to further understand these mechanisms.

It is unclear from the results of Experiment 2 whether participants were biased to repeat the strategy they used when first learning the task, or simply the strategy performed on the previous trial. We therefore conducted a third experiment in which we added a strategy reversal trial between the two phases. This aligns to the existing research investigating the number of trials needed to observe the perseveration effect and potential differences in the strength of the effect according to the number of preceding trials (Schillemans et al., 2012). This research suggests that a single application of a certain strategy can be sufficient to create a perseveration effect and bias participants' strategy choices on subsequent trials (Lemaire & Lecacheur, 2010; Schillemans et al., 2012). Consistent with this, experiment 3 showed that a single reversal-trial between phases 1 and 2 eliminated the group difference in the overall number of reminders set in phase 2 of the task. However, the single-trial analysis showed that the groups still differed in the temporal profile of strategy choices over the course of 10 trials. These findings suggest that a brief intervention, potentially involving just a single application of a particular strategy, could have a substantial effect on individuals' subsequent decisions whether to engage in cognitive offloading. However, the effect of such an intervention may fade over time.

The main limitations of this study relate to the online nature of the data collection and the reduced control over the setting in which participants provided their responses. We cannot exclude the possibility that some participants in Experiment 2 simply did not attend to the instructions explaining that they had a free choice whether to use reminders or their own memory in phase 2. Even though the instructions remained on the screen until the participant clicked a button to dismiss them, it is possible that some individuals did not fully understand them. However, this seems unlikely to fully explain the dramatic difference in phase-2 reminder setting between the two groups (47% versus 74%). Furthermore, even though the mean level of reminder setting did not differ between groups in Experiment 3, the groups did differ in the way that their phase 2 strategy shifted over time. This indicates that even when participants were required to try both strategies before being given a free choice, there was still a significant effect of the earlier strategy on subsequent choice, albeit a more subtle one than the large mean difference seen in Experiment 2.

Although online studies make it harder to control the setting in which participants take part, this can arguably increase their generalisability seeing as participation takes place at home rather than in an unfamiliar laboratory environment. The demographics of online samples also tend to be more representative of the general population, as was the case here with a larger age range than would be typical in experimental psychology. It is well established that age-related cognitive declines can affect memory performance, reducing accuracy, and impact upon the decision to use cognitive offloading behaviours (Weakley et al., 2019). Research has shown that individuals become more rigid in later adulthood (Lemaire et al., 2004; Lemaire & Lecacheur, 2001), and this rigidity might strengthen the bias towards the repeated application of a previously implemented strategy (Schillemans et al., 2012). This is particularly relevant when considering how age interacts with the effect of strategy history on cognitive offloading. Further delineating how age impacts upon the perseveration effect and on the decision to use a cognitive offloading strategy rather than relying on the internal memory represents an important direction for further research.

Another promising avenue for future research is the investigation of individual differences on offloading strategies. As pointed out by Boldt and Gilbert (Boldt & Gilbert, 2019), cognitive offloading is idiosyncratic in that different individuals use different aids to support their cognitive acts and different cognitive aids might be more appropriate than others in certain situations. We cannot exclude that some participants developed and adopted different strategies to solve the task. Future research should attempt to investigate individual differences in the use of cognitive offloading strategies and potential differences and similarities across different forms of offloading.

We note that although we have used the terminology of 'prospective memory', the task used here required participants to remember delayed intentions over a much shorter time-period than standard experimental tasks (e.g., Einstein & McDaniel, 1990; Graf & Uttl, 2001, for discussion; Smith, 2003). Nevertheless, there is evidence that the present task has significant external validity with respect to real-world prospective memory tasks conducted over a longer time scale: Gilbert (2015a) found that performance of the task administered here predicted participants' fulfilment of a naturalistic intention embedded within their everyday activities over the course of a week. Furthermore, the predictive ability of this task was greater than more standard event- and time-based prospective memory tasks. Nevertheless, it would be an interesting question for future research to investigate whether the results reported here hold in various experimental paradigms, involving a variety of retention intervals.

Conclusion

Technological advancement has profoundly changed how we access and store information. The possibility to offload into the environment more than we used to in the past offers the opportunity to reduce the burden of intention maintenance in internal memory. However, the consequences of this for future cognitive performance are not well understood. Our results suggest that short-term use of a cognitive offloading strategy has little influence on subsequent unaided performance, in a context where the information that was originally offloaded or stored internally is no longer relevant. However, there can be strong effects of prior offloading on subsequent strategy choice. Therefore, there is a risk that not engaging in cognitive offloading can impair subsequent strategy choice, in contexts where an offloading strategy is available and effective.

Acknowledgements: This work was supported by the Economic and Social Research Council (ESRC) under Grant ES/N018621/1.

Disclosure statement: The authors declare that they do not have competing interests.

References

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/10.1080/09602011.2011.582378

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265. https://doi.org/10.1037/0022-3514.74.5.1252

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*, *4*(1), 45. https://doi.org/10.1186/s41235-019-0195-y

Carr, N. G. (2008). *Is Google making us stupid?* The Atlantic.

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

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE*, *8*(3), e57410. https://doi.org/10.1371/journal.pone.0057410

Dang, J. (2018). An updated meta-analysis of the ego depletion effect. *Psychological Research*, *82*(4), 645–651. https://doi.org/10.1007/s00426-017-0862-x

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

Ellis, J. J., & Reingold, E. M. (2014). The Einstellung effect in anagram problem solving: evidence from eye movements. *Frontiers in Psychology*, *5*, 679. https://doi.org/10.3389/fpsyg.2014.00679

Ferguson, A. M., McLean, D., & Risko, E. F. (2015). Answers at your fingertips: Access to the Internet influences willingness to answer questions. *Consciousness and Cognition*, *37*, 91–102. https://doi.org/10.1016/j.concog.2015.08.008

Finley, J. R., Naaz, F., & Goh, F. W. (2018). *Memory and Technology: How We Use Information in the Brain and the World*. Springer, Cham.

Fisher, M., Goddu, M. K., & Keil, F. C. (2015). Searching for explanations: How the Internet inflates estimates of internal knowledge. *Journal of Experimental Psychology: General*, *144*(3), 674–687. https://doi.org/10.1037/xge0000070

Friese, M., Loschelder, D. D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2018). Is ego depletion real? An analysis of arguments. *Personality and Social Psychology Review*, 1–25. https://doi.org/10.1177/1088868318762183

Gilbert, S. J. (2015a). Strategic offloading of delayed intentions into the external environment. *Quarterly Journal of Experimental Psychology*, *68*(5), 971–992. https://doi.org/10.1080/17470218.2014.972963

Gilbert, S. J. (2015b). Strategic use of reminders: Influence of both domain-general and task-specific metacognitive confidence, independent of objective memory ability. *Consciousness and Cognition*, *33*, 245–260. https://doi.org/10.1016/j.concog.2015.01.006

Gilbert, S. J., Bird, A., Carpenter, J., Fleming, S. M., Sachdeva, C., & Tsai, P.-C. (2019). Optimal use of reminders: Metacognition, effort, and cognitive offloading. *Journal of Experimental Psychology: General*. https://doi.org/DOI: 10.1037/xge0000652

Graf, P., & Uttl, B. (2001). Prospective Memory: A New Focus for Research. *Consciousness and Cognition*, *10*(4), 437–450. https://doi.org/10.1006/ccog.2001.0504

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., Brand, R., Brandt, M. J., Brewer, G. A., Bruyneel, S., Calvillo, D. P., Campbell, W. K., Cannon, P. R., Carlucci, M., Carruth, N. P., Cheung, T., Crowell, A., De Ridder, D. T. D., Dewitte, S., … Zwienenberg, M. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, *11*(4), 546–573. https://doi.org/10.1177/1745691616652873

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

Herrmann, D., Brubaker, B., Yoder, C., Sheets, V., & Tio, A. (1999). Devices that remind. In F. Durso, R. Nickerson, R. Schvaneveldt, S. Dumais, D. Lindsay, & M. Chi (Eds.), *Handbook of applied cognition* (1st ed., pp. 377–407). Wiley.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social Psychology: Handbook of Basic Principles* (pp. 133–168). The Guilford Press.

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

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, *142*(2), 573–603. https://doi.org/10.1037/a0029146

Lemaire, P., Arnaud, L., & Lecacheur, M. (2004). Adults' age-related differences in adaptivity of strategy choices: Evidence from computational estimation. *Psychology and Aging*, *19*(3), 467–481. https://doi.org/10.1037/0882-7974.19.3.467

Lemaire, P., & Lecacheur, M. (2001). Older and younger adults' strategy use and execution in currency conversion tasks: insights from French franc to euro and euro to French franc conversions. *Journal of Experimental Psychology. Applied*, *7*(3), 195–206. http://www.ncbi.nlm.nih.gov/pubmed/11676098

Lemaire, P., & Lecacheur, M. (2010). Strategy switch costs in arithmetic problem solving. *Memory & Cognition*, *38*(3), 322–332. https://doi.org/10.3758/MC.38.3.322

Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs*, *54*(6), i–95. https://doi.org/10.1037/h0093502

Lurquin, J. H., & Miyake, A. (2017). Challenges to ego-depletion research go beyond the replication crisis: A need for tackling the conceptual crisis. *Frontiers in Psychology*, *8*, 568. https://doi.org/10.3389/fpsyg.2017.00568

Luwel, K., Schillemans, V., Onghena, P., & Verschaffel, L. (2009). Does switching between strategies within the same task involve a cost? *British Journal of Psychology*, *100*(4), 753–771. https://doi.org/10.1348/000712609X402801

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field*. Sage Publications. https://psycnet.apa.org/record/2007-00679-000

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

Plato. (1995). *Phaedrus*. (A. Nehamas & P. Woodruff, Trans.). Hackett Publishing Company. (Original work composed around 370 BC).

Quintana, D. S., & Williams, D. R. (2018). Bayesian alternatives for common null-hypothesis significance tests in psychiatry: a non-technical guide using JASP. *BMC Psychiatry*, *18*(1), 178. https://doi.org/10.1186/s12888-018-1761-4

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

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

Schillemans, V., Luwel, K., Bulté, I., Onghena, P., & Verschaffel, L. (2010). The influence of previous strategy use on individuals' subsequent strategy choice: Findings from a numerosity judgement task. *Psychologica Belgica*, *49*(4), 191. https://doi.org/10.5334/pb-49-4-191

Schillemans, V., Luwel, K., Ceulemans, E., Onghena, P., & Verschaffel, L. (2012). The effect of single versus repeated previous strategy use on individuals' subsequent strategy choice. *Psychologica Belgica*, *52*(4), 307. https://doi.org/10.5334/pb-52-4-307

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(3), 347–361. https://doi.org/10.1037/0278-7393.29.3.347

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

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. (2016). *Using the Internet to access information inflates future use of the Internet to access other information*. https://doi.org/10.1080/09658211.2016.1210171

Tamir, D. I., Templeton, E. M., Ward, A. F., & Zaki, J. (2018). Media usage diminishes memory for experiences. *Journal of Experimental Social Psychology*, *76*, 161–168. https://doi.org/10.1016/J.JESP.2018.01.006

Ward, A. F. (2013). *One with the Cloud: Why People Mistake the Internet's Knowledge for Their Own* [Doctoral Dissertation, Harvard University]. http://nrs.harvard.edu/urn-3:HUL.InstRepos:11004901

Weakley, A., & Schmitter-Edgecombe, M. (2019). Naturalistic assessment of task interruption in individuals with mild cognitive impairment. *Neuropsychology*, *33*(1), 1–12. https://doi.org/10.1037/neu0000481

Weakley, A., Weakley, A. T., & Schmitter-Edgecombe, M. (2019). Compensatory strategy use improves real-world functional performance in community dwelling older adults. *Neuropsychology*, *33*(8), 1121–1135. https://doi.org/10.1037/neu0000591

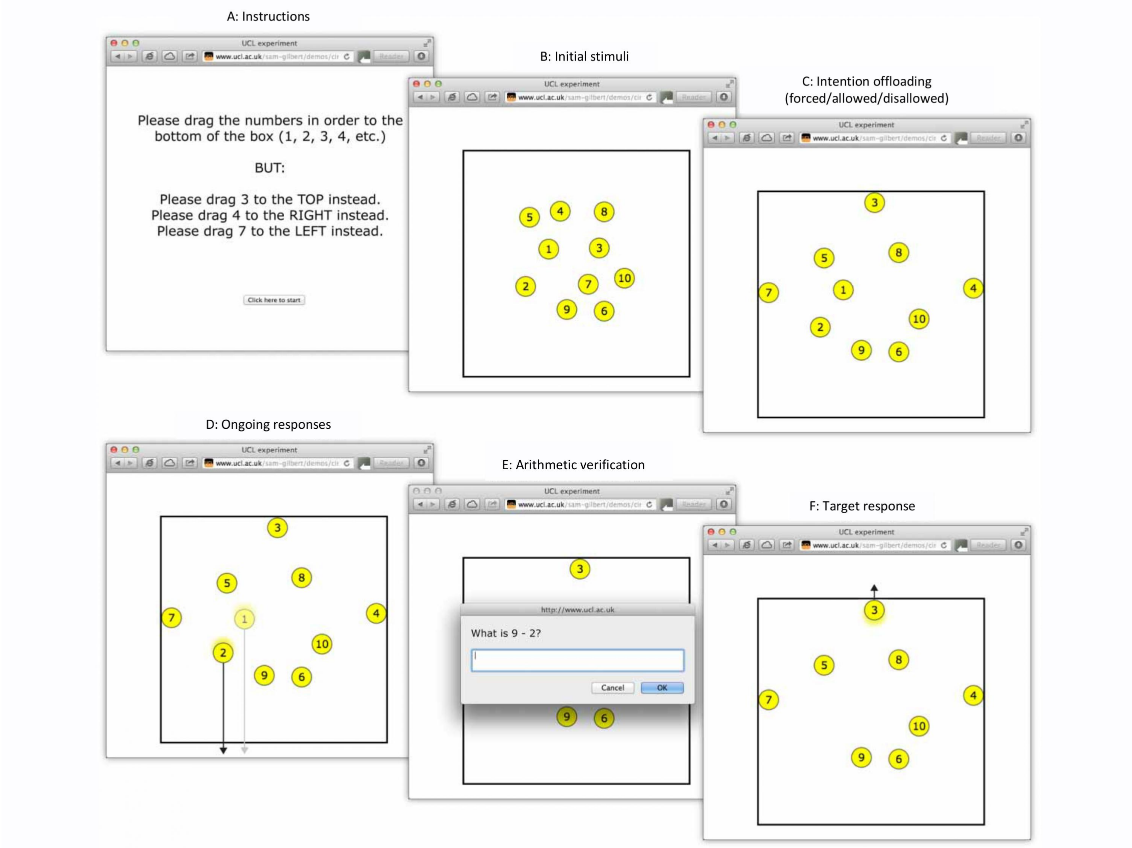

Figure 1. Schematic illustration of the intention-offloading task.

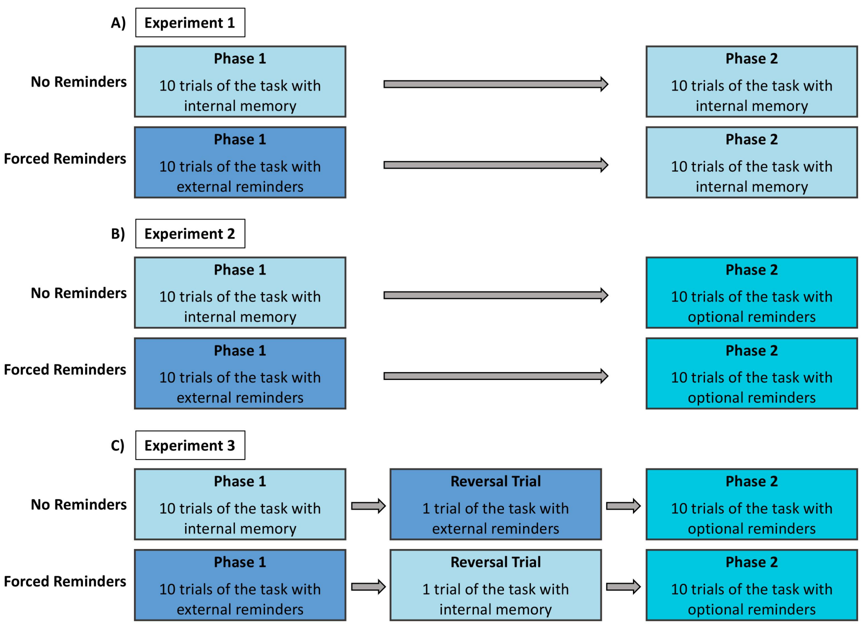

Figure 2. Experimental design for experiment 1 (A), experiment 2 (B) and experiment 3 (C).

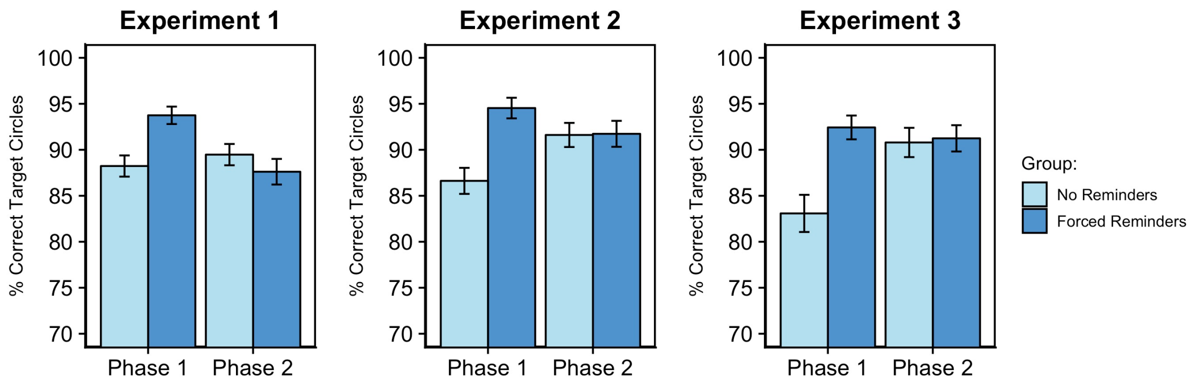

Figure 3. Mean target accuracy in experiments 1, 2, and 3. Error bars indicate standard error of the mean.

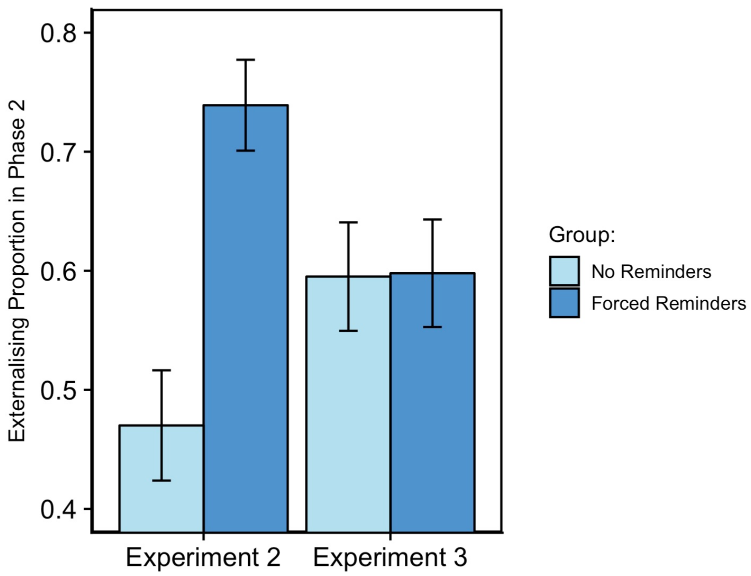

Figure 4. Mean externalising proportion in experiments 2 and 3. Error bars indicate standard error of the mean.

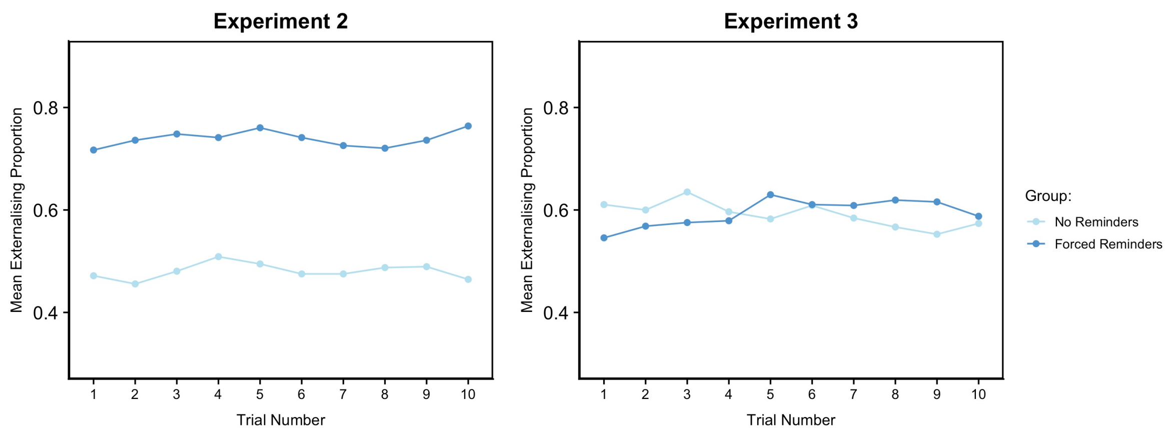

Figure 5. Mean externalising proportion across trials in phase 2 for experiment 2 and experiment 3.

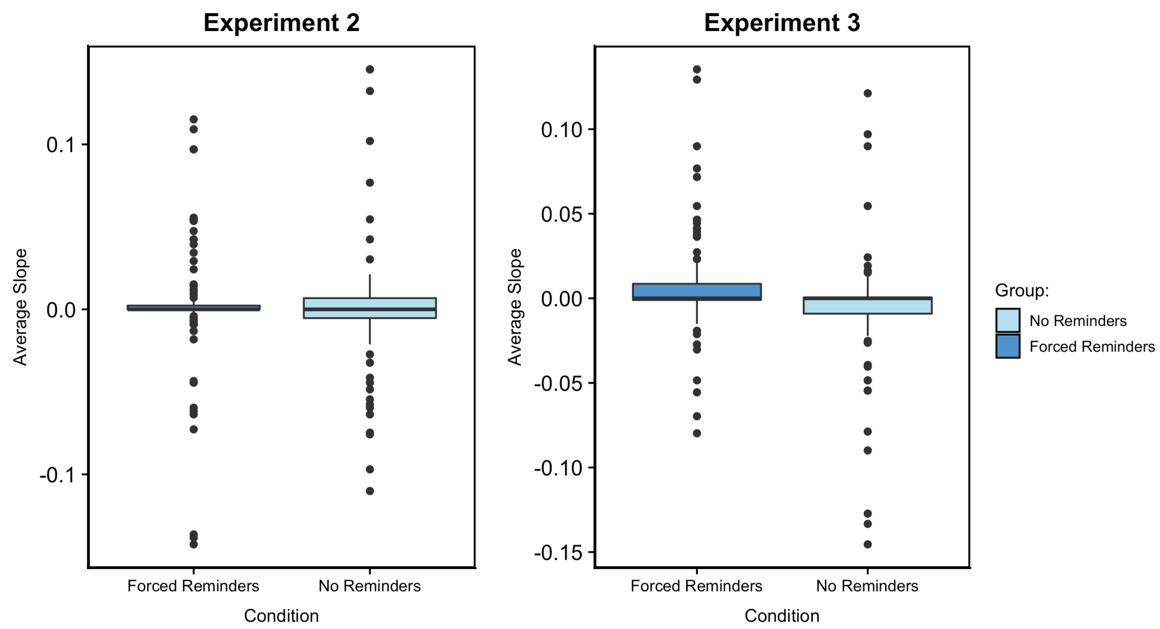

Figure 6. Group-wise distributions of individual slopes from a regression model in which trial predicts the externalising proportion for experiment 2 and experiment 3. The boxplots reflect the interquartile range (IQR) and the median. The whiskers span from the first quartile minus 1.5 times the IQR to the third quartile plus 1.5 times the IQR. Data points outside this inner fence are shown as individual data points.