Received: 31 October 2022 | Revised: 20 March 2023 | Accepted: 21 March 2023

DOI: 10.1002/acp.4067

SPECIAL ISSUE ARTICLE

WILEY

An investigation of the saving-enhanced memory effect: The role of test order and list saving

Pei-Chun Tsai[1] | Chhavi Sachdeva[1] | Sam J. Gilbert[1] | Chiara Scarampi[2,3]

[1]Institute of Cognitive Neuroscience, University College London, London, UK

[2]Centre for the Interdisciplinary Study of Gerontology and Vulnerability (CIGEV), University of Geneva, Geneva, Switzerland

[3]Swiss National Center of Competences in Research LIVES – Overcoming Vulnerability: Life Course Perspectives, Geneva and Lausanne, Switzerland

Correspondence
Pei-Chun Tsai, Institute of Cognitive Neuroscience, 17 Queen Square, London WC1N 3AR, UK.
Email: p.tsai.17@ucl.ac.uk

Funding information
Economic and Social Research Council, Grant/Award Number: ES/N018621/1

Abstract

Saving information onto external resources can improve memory for subsequent information—a phenomenon known as the saving-enhanced memory effect. This article reports two preregistered online experiments investigating (A) whether this effect holds when to-be-remembered information is presented before the saved information and (B) whether people choose the most advantageous strategy when given free choice of which information to save. Participants studied two lists of words; test order and whether and which list was saved (and re-presented again later) were manipulated. The saving-enhanced memory effect was only found when the first list (List A) was saved and tested after the second list (List B). When free to choose which list to save, participants preferred to save List A, but only when it was recalled after List B—that is, when it benefited memory. These findings suggest boundary conditions for the saving-enhanced memory effect and that people offload the most profitable information.

KEYWORDS
cognitive offloading, memory, metacognition, selectivity

1 | INTRODUCTION

Technology is changing the world by enabling greater access to information and overcoming the limitations of the human mind (Miller, 1956) through the use of external devices such as smartphones and computers. The use of such physical actions to simplify the cognitive demands of a task is known as "cognitive offloading" (Risko & Gilbert, 2016). Offloading large amounts of information onto external memory stores helps us cope with the increasing volume of information that we are required to process daily. Being able to temporarily forget saved information that is irrelevant at present has the fundamental advantage of reducing memory load (Herrmann et al., 1999). Furthermore, it enables us to focus our limited cognitive resources on current tasks while temporarily preserving non-essential information in an external store for later use. The integration of digital devices into our lives raises relevant questions on how interacting with these devices impacts and alters our memory processes. The current study examined the impact of saving information on digital devices on the memory of both previously and newly encoded information.

Previous research has shown that using digital devices as external, extended memory stores to offload temporarily-irrelevant information can enhance subsequent memory performance (Storm & Stone, 2015), a phenomenon referred to as the *saving-enhanced memory* effect. Storm and Stone (2015) conducted an experiment where participants were presented with two lists of eight words and were subsequently tested on their memory for both lists. On some trials, participants were allowed to "save" the first list, while on other trials they were forced to remember both lists using unaided memory. The authors found that when the contents of a file were saved and restudied before test,

Pei-Chun Tsai and Chhavi Sachdeva contributed equally to this work.

The preregistration of the study and the data are accessible at https://osf.io/vb8te/ (Experiment 1) and https://osf.io/m9ths/ (Experiment 2). The experiment materials are available at https://app.gorilla.sc/openmaterials/220207.

participants could remember a higher proportion of information from that file than when it was not saved. Furthermore, saving a file before studying a new file significantly improved recall of the contents of the new file suggesting that saving facilitates the encoding and remembering of new information. In a further study, Runge et al. (2020) showed that the saving-enhanced memory effect holds (and is even stronger) when using motor sequences instead of word lists. Additionally, in an adaptation of Storm and Stone (2015)'s paradigm, Runge et al. (2019) demonstrated that saving a studied list not only improved recall of a subsequently learnt word list but also improved performance in a subsequent modular arithmetic task. This suggests that the benefits of offloading memory onto external sources are not limited to memory performance but extend to improving performance in a subsequent unrelated task—that is, a *saving-enhanced performance* effect.

Speculations have been made about the mechanisms underlying the saving-enhanced memory effect which parallel the phenomena of list method directed forgetting (LMDF; Runge et al., 2020, 2021), where cueing participants to forget a previously studied list (List A) and remember a new list instead (List B) also leads to the forgetting of the first list and enhanced memory of the second (see Bäuml et al., 2010; Sahakyan et al., 2013 for a review). Previous studies in the LMDF field have documented various processes that contribute to the memory enhancement effect. These include selective rehearsal of List 2 (Bjork, 1970), reduced proactive interference (Bjork & Bjork, 1996), switching encoding strategies (Sahakyan & Delaney, 2003), a change in participants' internal context (Sahakyan & Kelley, 2002), and reset of encoding (Pastötter & Bäuml, 2010). Runge et al. (2020) conducted a study to examine the theoretical explanations for the saving-enhanced memory effect and found evidence for the enhanced encoding explanation (Pastötter & Bäuml, 2010). According to this account, saving information reduces memory load, freeing cognitive resources that can be allocated to subsequent tasks. The authors also replicated the cost effects of cognitive offloading documented by previous research in the field (e.g., Henkel, 2014; Sparrow et al., 2011), where saving information externally reduced the accessibility of the saved information (Runge et al., 2020). This finding supports the reduced interference at recall hypothesis (Bjork & Bjork, 1996)—that is, saved information can be temporarily forgotten, reducing proactive interference on subsequently encoded information.

The findings reviewed above showcase the benefits of cognitive offloading, where using technology to supplement our memory allows us to expand our memory capacity. However, one limitation of these studies is that they investigated the benefits of saving information onto an external store in a very specific paradigm, where the events always follow the same sequence: a first list of words is saved, a second list is studied, then the second list is tested, and finally the first list is restudied and tested. However, in our everyday life, we are usually confronted with multiple tasks that are active at the same time and we are constantly required to manage information. For example, we may have already encoded some information in internal memory and then later decide to rely on an external store for subsequent information. Therefore, the first aim of the current study was to investigate if 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).

Research in LMDF has provided contradicting results when participants were instructed to forget the second list. Sahakyan (2004) found that attempting to forget List B had both direct and indirect impacts on forgetting (also see Racsmány et al., 2019). In their experiment, Sahakyan (2004) presented participants with three lists of words that 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 this list 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 (visually and auditorily) and the discriminability between the two lists (relevant vs. irrelevant information). This result was subsequently replicated with both short and long lists (Kliegl et al., 2020).

The contrasting results between the two experiments could be explained by one difference between their 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 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 causing 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). To further examine this effect in the context of cognitive offloading, a second aim of the current study 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 this order was reversed for the other half of the participants. This is again in line with the notion that in everyday life we manage different tasks simultaneously, and the order in which different pieces of information are encoded and retrieved varies continuously. The manipulation of which list is encoded and which list is tested first will provide further insights on the theoretical accounts of the saving-enhanced memory effect.

Exploring the mechanisms underlying the saving-enhanced memory effect is important. Equally important is gaining a deeper understanding of the processes underlying people's decisions to offload. Previous research in the cognitive offloading field has shown that the decision to offload our cognition onto the environment is triggered by some form of metacognitive processing (Dunn & Risko, 2016; Gilbert, 2015). Metacognition involves both a monitoring mechanism based on the subjective assessment of one's own cognitive processing and a control mechanism, which involves the implementation of strategies to regulate cognition (Flavell, 1979; Nelson & Narens, 1990). Cognitive offloading can be seen as a form of metacognitive control based on the evaluation of one's own cognitive abilities, and the estimation of the costs and benefits associated with reliance on external strategies versus internal processing (Risko & Dunn, 2015). A well-established result in the field is that individuals who have low confidence in their ability to remember information are more likely to store it externally (see e.g., Boldt & Gilbert, 2019; Gilbert, 2015; Risko & Dunn, 2015). However, the literature so far has focused on comparing reliance on internal memory versus external stores. How individuals choose between different offloading strategies remains to be established. Accordingly, in the current study, we were interested in investigating not only the role of metacognition in deciding whether or not to save temporary non-critical information onto external memory stores but also which information is saved. That is, while Experiment 1 established the conditions under which saving information leads to the best memory performance, Experiment 2 investigated the role of participants' metacognitive beliefs on saving decisions and examined whether individuals implement effective metacognitive strategies by relying on external memory aids when these are most advantageous.

2 | EXPERIMENT 1

The aims of this experiment were twofold. The first was to compare the benefits of offloading List A versus the benefits of 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 words. They were then subsequently tested on both lists separately. We hypothesized that (*H*1) if saving a file works as a forget cue, we would find increased recall for List B-items on *List A-Saving* trials and spared or reduced recall for List A-items on *List B-Saving* trials.

The second aim was to investigate whether the testing order of Lists A and B influenced the saving-enhanced memory effect of List B when List A was saved. Half of the participants were asked to recall List A-items before List B and this test order was reversed for the other half of participants, who were asked to recall List B-items before List A. We predicted that (*H*2) on *List A-Saving* trials, the saving-enhanced memory effect for List B would be larger when List B was recalled first than when List A was recalled first, as found for forget cues in LMDF literature. Furthermore, we predicted that (*H*3) on *List A-Saving* trials, the proportion of words recalled from List A would be larger when it was recalled second than when it was recalled first, as List B would already have been tested and participants would be able to completely focus on List A when restudying it.

Before commencing data collection, we preregistered our hypotheses, experimental procedure and analysis plan (https://osf.io/vb8te/).

2.1 | Method

2.1.1 | Participants

A total of 102 participants (51 in each group; *mean* age = 35.51 years; SD age = 11.82 years; 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 by UCL Research Ethics Committee (1584/003), and participants provided informed consent before participating in the study. Participation took approximately 40 min, 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). Our power calculation was based on the results of Storm and Stone (2015 Experiment 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, we used a more conservative approach where we halved this number (.93). This resulted in an effect size (d_z) of .465. To achieve 90% power to replicate an effect of this size (two-tailed test, $\alpha = .05$), a sample of 51 participants was required. Since the test order factor in this experiment was manipulated between subjects, the total sample was 102 participants, with 51 participants in each test order group.

Participants whose memory performance (averaged across conditions) exceeded three median absolute deviation units (MAD; Leys et al., 2013) (outliers) were excluded ($n = 8$) and replaced so that the sample size totaled 102 participants. Furthermore, three participants reported cheating (e.g., writing things down or taking pictures) and were thus excluded and replaced.

2.1.2 | Design

This task was programmed using Gorilla (https://gorilla.sc; Anwyl-Irvine et al., 2020) and had two manipulations. The first was whether and which list was saved and restudied. This was manipulated within subjects so that 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 their unaided memory where neither list could be saved (*No-Saving trials*).

The second manipulation was 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 1 for a schematic representation of the task).

2.1.3 | Materials

For this experiment, 192 common nouns (4–7 letters long) were selected from the Paivio Word List Generator (http://euclid.psych.yorku.ca/shiny/Paivio/). For each participant, 24 lists of eight words were randomly created. Half of these lists (i.e., 12) were assigned as List A while the other half were assigned as List B. During the break between trials, 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.

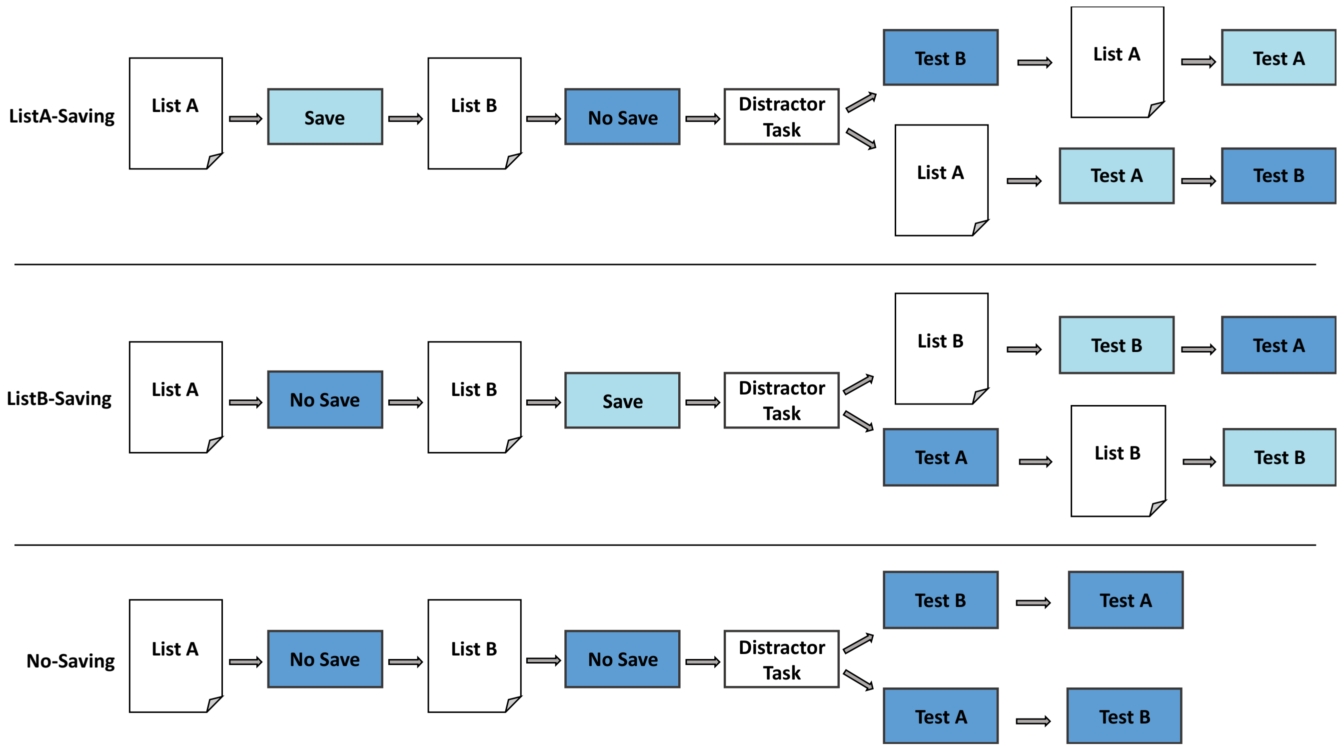

FIGURE 1 Schematic illustration of the trial sequence in the different conditions. Participants were instructed to study List A and List B. They performed four trials in each of the three conditions detailed in the figure: On *List A-Saving trials*, participants saved List A before studying List B; on *List B-Saving trials*, participants saved list B before the distractor task; on *No-Saving 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, whereas 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.

2.1.4 | Procedure

Participants provided informed consent and were randomly assigned to one of the test order conditions. They were then informed that each trial of the task would involve studying the contents of two lists and subsequently being tested on the contents of those lists. They were asked to recall as many words as possible. Participants completed three practice trials where they were introduced to each of the three conditions (*No-Saving*, *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 presented with a short comprehension quiz that tested their knowledge of the different elements of the task. For every mistake they made, participants received further clarification on what that task component 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 were *List B-saving trials*, and the remaining third were *No-Saving trials*. These trials were presented in randomized order. On each trial, participants studied List A and List B for 15 s each. On *List A-Saving trials*, participants were instructed to save List A after studying it by actively pressing a button labeled "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 labeled "Save" on the screen. On *No-Saving trials*, participants were not allowed to save or restudy either list. Participants were informed that saving a list would ensure that they would restudy it prior to test. That is, although participants could not view the saved list while it was being tested, they were able to restudy it just before it was tested.

After the study phase, there was a short 20-s delay during which participants were asked to count backward 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 they 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 s to type all the words they could recall from the instructed list. On *List A-Saving trials*, before the recall test for List A, participants were prompted to press a button to reopen the saved list. They then had 15 s to restudy List A before being tested on it. Similarly, on *List B-Saving trials*, before the recall test for List B, participants were prompted to press a button to reopen the saved list. They then had

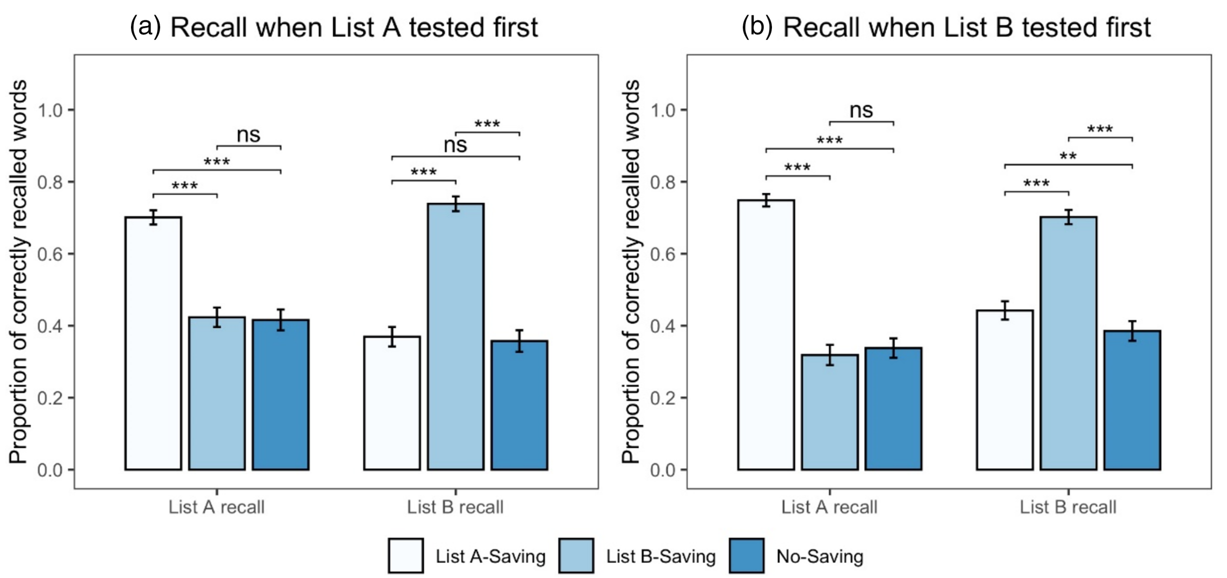

FIGURE 2 Recall of List A and List B as a function of saving condition and test order in Experiment 1. Error bars represent standard errors of the mean.

15 s to restudy List B before being tested on it. On *No-Saving trials*, participants recalled the two lists without restudying either list. After each trial, participants were given a "Spot-the-difference" puzzle as a distractor task for 1 min before beginning the subsequent trial. The timings for each experiment phase were the same for every participant.

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).

2.2 | Results

Analyses were conducted using R (version 4.0.3). Bayes factors were calculated using JASP (version 0.14.1) to assess evidence for null effects. In accordance with Jeffreys (1961), the Bayes factor scores were interpreted as follows: a score between 1–3 was interpreted as anecdotal, 3–10 as substantial, 10–30 as strong, 30–100 as very strong, and 100+ as extreme. Analyses were conducted as per our pre-registered plan. Any additional analyses conducted are highlighted. All analyses were two-tailed with $\alpha = .05$.

2.2.1 | Recall performance for List B

We first investigated recall performance for List B to examine whether the saving-enhanced memory effect from Storm and Stone (2015) was replicated. To investigate the proportion of words recalled from List B, a 3 × 2 mixed ANOVA was conducted with within-subjects factor saving condition (List A-Saving vs. List B-Saving vs. No-Saving) and between-subjects factor test order (List A → List B vs. List B → List A) was conducted. No significant main effect of test order was found ($F(1, 100) = 0.54$, $p = .465$, $\eta^2_p = .005$, $BF_{01} = 5.35$) but a significant main effect of saving condition ($F(1, 100) = 230.35$, $p < .001$, $\eta^2_p = .697$, $BF_{10} = 1.171e^{+49}$) and a significant interaction ($F(1, 100) = 4.73$, $p = .014$, $\eta^2_p = .045$, $BF_{10} = 1.198e^{+49}$) were found.

Three follow-up paired *t*-tests[1] were computed separately for the two test order groups. In the group where List B was tested first (see Figure 2b), we replicated the result of Storm and Stone (2015), where there was a significant difference between List A-Saving trials and No-Saving trials ($t(50) = 3.27$, $p = .002$, $d = .46$, $BF_{10} = 15.72$). In these trials, List B-recall was higher when List A was saved than when neither list was saved. Furthermore, there was a significant difference between List A-Saving and List B-Saving trials ($t(50) = 10.28$, $p < .001$, $d = 1.44$, $BF_{10} = 1.180e^{+11}$) where List B-recall was higher on List B-saving trials than on List A-Saving trials. There was also a significant difference between List B-Saving trials and No-Saving trials ($t(50) = 12.69$, $p < .001$, $d = 1.78$, $BF_{10} = 1.946e^{+14}$) where List B-recall was higher on List B-Saving trials than on No-Saving trials.

In the group where List A was tested first (see Figure 2a), there was a significant difference between List A-Saving and List B-Saving trials ($t(50) = 14.26$, $p < .001$, $d = 2.00$, $BF_{10} = 1.682e^{+16}$) and No-Saving and List B-Saving trials ($t(50) = 11.53$, $p < .001$, $d = 1.61$, $BF_{10} = 6.117e^{+12}$) where recall of List B was higher on List B-Saving trials than on List A-Saving or No-Saving trials. However, there was no difference between the List A-Saving and No-Saving trials ($t(50) = 0.53$, $p = .595$, $d = .07$, $BF_{01} = 5.56$). These results suggest that saving List A only improved recall of List B when List B was tested first.

Three additional follow-up independent samples *t*-tests were conducted separately in the three saving conditions to compare List B-recall when it was tested first compared to when it was tested second. On List A-Saving trials, recall of List B was slightly higher when it was tested first ($M = .44$, $SD = .18$) than when it was tested second ($M = .37$, $SD = .20$), but this result did not reach the conventional threshold for significance ($t(99.62) = 1.96$, $p = .053$, $d = .39$, $BF_{10} = 1.14$) and Bayes Factor found only anecdotal evidence in favor of the alternative hypothesis. On No-Saving (List A → List B: $M = .39$, $SD = .20$; List B → List A: $M = .36$, $SD = .21$; $t(50) = 0.69$, $p = .500$, $d = .14$, $BF_{01} = 3.85$) and List B-Saving trials (List A → List B: $M = .70$, $SD = .14$; List B → List A: $M = .74$, $SD = .15$; $t(50) = 1.29$, $p = .199$, $d = .26$, $BF_{01} = 2.27$), List B recall was not affected by test order.

2.2.2 | Recall performance for List A

To investigate the proportion of words correctly recalled from List A, a 3×2 mixed ANOVA analogous to the one above was conducted with within-subjects factor saving condition (List A-Saving vs. List B-Saving vs. No-Saving) and between-subjects factor test order (List A → List B vs. List B → List A). Similar to the results of List B, the main effect of test order was not significant ($F(1, 100) = 2.35$, $p = .129$, $\eta^2_p = .023$, $BF_{01} = 2.70$), but the main effect of saving condition ($F(1, 100) = 280.46$, $p < .001$, $\eta^2_p = .737$, $BF_{10} = 1.917e^{+53}$) and the interaction ($F(1, 100) = 11.36$, $p < .001$, $\eta^2_p = .102$, $BF_{10} = 1.122e^{+56}$) were significant.

Three follow-up paired t-tests were computed separately for the two test order groups. In the group where List A was tested first (see Figure 2a), there was a significant difference between No-Saving and List A-Saving trials ($t(50) = 10.55$, $p < .001$, $d = 1.48$, $BF_{10} = 2.863e^{+11}$), and between List A-Saving and List B-Saving trials ($t(50) = 12.34$, $p < .001$, $d = 1.73$, $BF_{10} = 6.980e^{+13}$) where List A-recall was higher on List A-Saving trials than on No-Saving trials or List B-Saving trials. There was no significant difference in the proportion of List A-words recalled on No-Saving and List B-Saving trials ($t(50) = 0.41$, $p = .685$, $d = 0.06$, $BF_{01} = 6.06$), 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 (see Figure 2b), there was a significant difference between No-Saving and List A-Saving trials ($t(50) = 155$, $p < .001$, $d = 2.1$, $BF_{10} = 1.237e^{+17}$) and List A-Saving and List B-Saving trials ($t(50) = 13.81$, $p < .001$, $d = 1.93$, $BF_{10} = 4.784e^{+15}$). Recall of List A was higher on List A-Saving trials than on List B-Saving and No-Saving trials. Once again, no significant difference was found between List B-Saving and No-Saving trials ($t(50) = 1.25$, $p = .218$, $d = 0.18$, $BF_{01} = 3.16$).

Three additional follow-up independent-samples t-tests were conducted separately in the three saving conditions to compare List A-recall when it was tested first compared to when it was tested second. On List A-Saving trials, recall of List A did not differ significantly when it was tested second ($M = .75$, $SD = .12$) and when it was tested first ($M = .70$, $SD = .21$; $t(98.43) = 1.83$, $p = .07$, $d = .36$, $BF_{01} = 1.09$). In contrast, List A-recall was significantly higher when it was tested first on List B-Saving trials (List A → List B: $M = .42$, $SD = .19$; List B → List A: $M = .32$, $SD = .20$; $t(99.76) = 2.69$, $p = .008$, $d = .53$, $BF_{10} = 4.88$) and a marginally-significant effect was seen on No-Saving trials (List A → List B: $M = .42$, $SD = .21$; List B → List A: $M = .34$, $SD = .19$; $t(99.45) = 1.98$, $p = .051$, $d = .39$, $BF_{10} = 1.17$; see Figure 2a).

2.2.3 | Recall performance for the list that was saved

To investigate recall performance for the list that was saved—that is, recall performance of List A when List A was saved and recall performance of List B when List B was saved—, a 2×2 mixed ANOVA with within-subjects factor saving condition (List A-Saving vs. List B-Saving) and between-subjects factor test order (List A → List B vs. List B → List A) was conducted on the proportion of correctly recalled words of the list that was saved and restudied. There was no significant main effect of saving condition ($F(1, 100) = 0.14$, $p = .709$, $\eta_p^2 = .001$, $BF_{01} = 6.37$) and test order ($F(1, 100) = 0.05$, $p = .824$, $\eta_p^2 < .001$, $BF_{01} = 3.46$) but there was a significant interaction ($F(1, 100) = 13.62$, $p < .001$, $\eta_p^2 = .12$, $BF_{10} = 3.37$).

To qualify our interaction, we conducted follow-up paired t-tests 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 ($M = .70$, $SD = .14$) ($t(50) = -2.16$, $p = .035$, $d = .30$, $BF_{10} = 1.29$).

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 ($M = .70$, $SD = .14$) ($t(50) = 3.16$, $p = .003$, $d = .44$, $BF_{10} = 11.81$). Therefore, in both groups, performance was better for whichever list was restudied and tested second.

2.2.4 | Recall performance for the list that was not saved

To investigate recall performance for the list that was not saved—that is, recall performance of List A when List B was saved and recall performance of List B when List A was saved—, a 2×2 mixed ANOVA with within-subjects factor saving condition (List A-Saving vs. List B-Saving) and between-subjects factor test order (List A → List B vs. List B → List A) was conducted on the proportion of correctly recalled words of the list that was not saved. There was no significant main effect of test order ($F(1, 100) = 0.22$, $p = .644$, $\eta_p^2 = .002$, $BF_{01} = 3.61$), but there was a significant main effect of saving condition ($F(1, 100) = 4.45$, $p = .037$, $\eta_p^2 = .043$, $BF_{01} = 1.35$) 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 ($M = .37$, $SD = .20$). We also found a significant interaction ($F(1, 100) = 28.81$, $p < .001$, $\eta_p^2 = .22$, $BF_{10} = 5734.32$).

To qualify the interaction, we conducted follow-up paired t-tests 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 ($M = .37$, $SD = .19$) ($t(50) = 2.49$, $p = .016$, $d = .35$, $BF_{10} = 2.48$).

Conversely, in the group where List B was tested first, the proportion of words 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 ($M = .32$, $SD = .20$) ($t(50) = 4.94$, $p < .001$, $d = .69$, $BF_{10} = 2084.85$).

2.3 | 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 our first hypothesis, we found a saving-enhanced memory effect for List B when List A was saved but no benefit for List A-items when List B was saved. 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 instructed to forget List B. This held 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 our second hypothesis, where we 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. We did not find support for the third hypothesis, where we predicted that there would be an effect of testing order on List A-Saving trials. However, we 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 is true for the group where List B was tested first.

3 | EXPERIMENT 2

The results of Experiment 1 suggest that when presented with two memory tasks and allowed to offload one of the two, it is advantageous to offload the first task and rely on internal memory for the second task. As a follow-up, the current study aimed at extending the findings of Experiment 1 by investigating whether, when given free choice, participants would have a preference towards saving a specific list. More precisely, we tested whether participants would choose to save and restudy the most effective list, that is, List A. We did this by adapting the paradigm used in Experiment 1. At the beginning of each trial, participants decided between saving a list and not saving a list. If they chose to save a list, they decided between saving List A and saving List B. In line with Experiment 1, we also investigated whether participants' preferences were influenced by the order in which the two tests were recalled. We predicted that (*HP*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, (*HP*2) they would prefer to save List A when List B was tested first, and they would prefer to save List B when List A was tested first.

Previous literature has shown that decisions to offload information onto an external store are influenced by confidence in one's cognitive abilities (e.g., individuals are more likely to set reminders when they have low confidence in their memory abilities; Boldt & Gilbert, 2019; Gilbert, 2015; Gilbert et al., 2020). Accordingly, one of the aims of this experiment was to study whether participants' preferences would be related to their confidence in their ability to remember the list that was not saved. Therefore, we asked participants to estimate their recall for List-A and List-B items in the different experimental conditions both before starting the task and at the end of the main task. We hypothesized that (*HP*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. That is, the less confident they were, the more they would choose to save a list. Furthermore, we anticipated that (*HP*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 confidence for List B-recall when List A was saved and List A-recall when List B was saved.

3.1 | Method

3.1.1 | Participants

A total of 88 participants (*mean* age = 39.6 years; SD age = 12.9 years; 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 min, 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). We decided to power the experiment to detect a medium effect size ($d = 0.5$) in 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, $\alpha = .05$), 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.

3.1.2 | Design

This experiment was programmed using Gorilla (https://gorilla.sc/) and contained 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 or no list. If they chose to save one 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 the test phase. On the other 50% of the save trials, participants were not re-presented with the list and so could not restudy the list they had chosen to save before the test phase. This was done to ensure that participants encoded both lists even if only one was saved. This made the design of the current experiment comparable to 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.

3.1.3 | Materials

This experiment used the same stimuli materials as Experiment 1.

3.1.4 | Procedure

The procedure differed from Experiment 1 in that at the beginning of each of the 12 trials, participants were given a choice to either save List A, save List B, or perform the task using their own memory. Participants were informed that on save trials, they would be allowed to restudy the saved list on only half of the trials. Participants' pre-task and post-task confidence ratings were also collected. For these ratings, we asked participants to estimate their recall of (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 B when List A was saved, (5) List A when List B was saved, and finally, and (6) List B when List B was saved (for further details on the procedure see the pre-registration protocol at https://osf.io/gvtd4).

3.2 | Results

Analyses were conducted as per our pre-registered plan. We also reported some additional exploratory analyses conducted on participants' confidence ratings.

We first investigated whether participants preferred saving a list over using their unaided memory. To investigate this, we computed the proportion of times participants chose to save a list and then conducted a one-sample t-test against $\frac{1}{2}$, seeing as participants could only choose one of two options (Save vs. No-Save: these measures are inverse of each other). We found that participants chose to save a list significantly more than half the time ($t(87) = 3.02$, $p = .003$, $d = .32$, $BF_{10} = 8.04$), showing a preference towards saving a list ($M = .63$, SD $= .41$) rather than using unaided memory ($M = .37$, SD $= .41$; see Figure 3a).

We then examined participants' preferences to offload a specific list. In other words, we wanted to examine whether participants preferred offloading List A or List B. For this analysis, we excluded all trials in which participants chose not to save a list. We then computed the proportion of times participants chose to save List A over List B (these measures are inverse of each other) and conducted a one-sample t-test against $\frac{1}{2}$. We found that participants chose to save List A significantly more than half the time ($t(74) = 3.08$, $p = .003$, $d = .36$, $BF_{10} = 9.62$), where participants showed a preference to save List A ($M = .63$, SD $= .37$) rather than saving List B ($M = .37$, SD $= .37$; see Figure 3b).

Next, two independent-samples t-tests were conducted. The first investigated if participants' preference towards saving a list versus using unaided memory depended on the order in which the two lists were recalled. The proportion of times participants chose to offload a list rather than relying on internal memory did not differ significantly in the two test order conditions (List A → List B: $M = .58$, SD $= .40$; List B → List A: $M = .69$, SD $= .42$; $t(85.88) = 1.27$, $p = .21$, $d = .208$, $BF_{01} = 2.22$). These results did not change when controlling for participants' metacognitive predictions on their ability to remember List-A and List-B items using unaided memory only, in an ANCOVA model[2] ($BF_{01} = 1.00$). However, when controlling for metacognitive post-task ratings in an ANCOVA model instead, Bayes Factors found anecdotal evidence in favor of the model including only post-task confidence for List A-recall when neither list was saved ($BF_{10} = 1.46$).

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 List A-saving was significantly larger when List B was tested first ($M = .76$, SD $= .31$) than when List A was tested first ($M = .51$, SD $= .39$; $t(70.52) = 3.01$, $p = .004$, $d = .69$, $BF_{10} = 10.31$; see Figure 3). Controlling for participants' metacognitive predictions regarding their ability to recall List A- and List B-items using unaided memory, did not alter the results in an ANCOVA. The Bayes Factors provided strong support for a model that only included test order as a factor ($BF_{10} = 10.31$). When controlling for metacognitive post-task ratings instead, the main effect of test order ($F(1, 71) = 5.67$, $p = .02$, $\eta^2_p = .074$, $BF_{10} = 10.31$) and post-task confidence for List A-recall when neither list was saved ($F(1, 71) = 4.01$, $p = .049$, $\eta^2_p = .053$, $BF_{10} = .48$) were both significant. Post-task confidence for List B-recall when neither list was saved was not significant ($F(1, 71) = 3.08$, $p = .083$, $\eta^2_p = .042$, $BF_{01} = 4.13$).

The same results held when the difference in predicted 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) was added as a covariate in an ANCOVA model. Bayes Factors strongly supported the model that only included test order as a factor ($BF_{10} = 10.31$). When controlling for the difference in post-task accuracy ratings instead, both test order ($F(1, 72) = 5.05$, $p = .028$, $\eta^2_p = .066$, $BF_{10} = 10.31$) and the difference in post-task

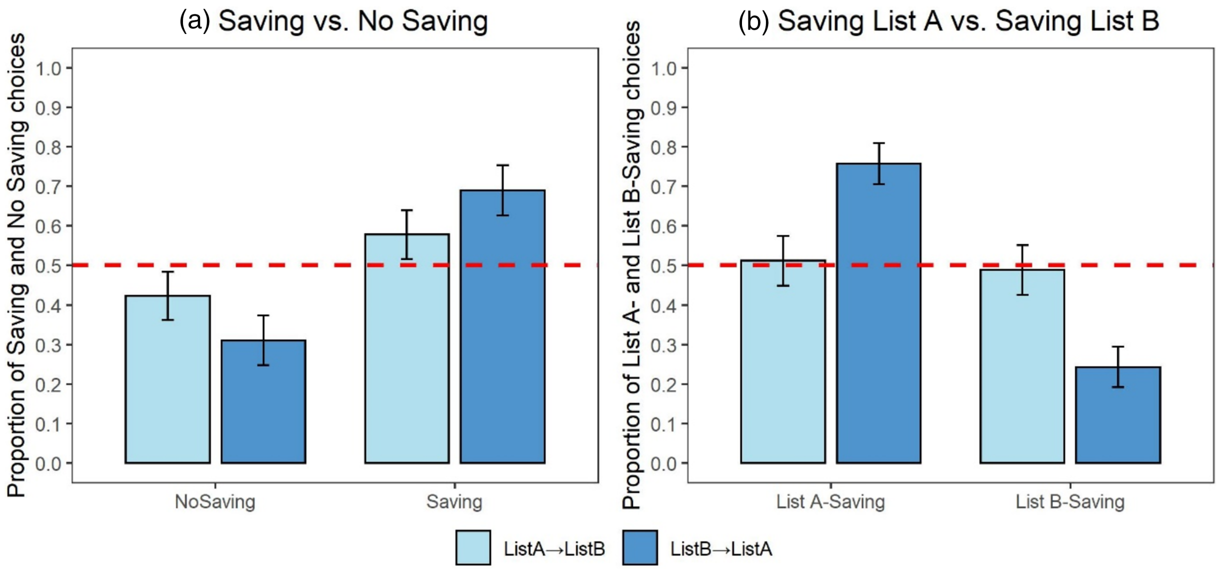

FIGURE 3 Participants' offloading choices in Experiment 2. This figure represents the proportion of times participants chose to (a) save a list versus not save a list and (b) save List A versus save List B, as a function of test order. The red dotted line denotes chance level at 0.5. Error bars represent standard errors of the mean. The proportion of times participants chose to save a list was calculated as the number of trials on which participants chose to save a list divided by the total number of trials. The proportion of times participants chose not to save a list is the inverse of this measure. The proportion of times participants chose to save List A was calculated as the number of trials on which participants chose to save List A divided by the total number of trials. The proportion of times participants chose to save List B is an inverse of this measure.

confidence (F(1, 72) = 5.72, p = .019, η^2_p = .074, BF_{10} = 13.67) were significant.

We also ran two one-sample t-tests against $\frac{1}{2}$ on the proportion of times participants chose to save List A separately in the two test order conditions. We found that participants had a preference towards offloading List A only when List B was tested first (List A → List B: t(37) = .18, p = .86, d = .03, BF_{01} = 5.64; List B → List A: t(36) = 4.99, p < .001, d = .82, BF_{10} = 1369.94).

A correlational analysis was conducted separately for List A and List B to investigate if the proportion of times a list was saved was associated with participants' confidence in their ability to recall List A or List B using unaided memory only. The proportion of times participants chose to save List A did not correlate with predicted accuracy on List A-recall when performing the task using unaided memory (r(73) = −.15, p = .192, BF_{01} = 3.03), nor with post-task confidence rating on List A unaided recall (r(73) = −.15, p = .209, BF_{01} = 3.23). Similarly, the proportion of times participants chose to save List B did not correlate with predicted accuracy on List B recall when performing the task using unaided memory (r(73) = .08, p = .501, BF_{01} = 5.56), nor with post-task confidence judgment on List B unaided recall (r(73) = −.02, p = .857, BF_{01} = 6.67).

In addition to the analyses above, we correlated the proportion of times participants saved List A 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 = .582, BF_{01} = 5.88).

The same set of correlations 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, BF_{10} = 12.35), where List A (List B) saving increased as the difference between the two post-task confidence scores increased (see Figure 4).

3.3 | Discussion

Overall, the results of the present experiment found that when given a choice, participants preferred to save a list rather than use unaided memory. This was in line with our first hypothesis. We only found partial support for our second hypothesis, where 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 allowed 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, that is, List A (at least in this paradigm). We did not find support for our third hypothesis as neither the pre-task nor post-task confidence ratings of recalling List A or List B using unaided memory were correlated with the preference for saving. We did, however, find partial support for the fourth hypothesis 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.

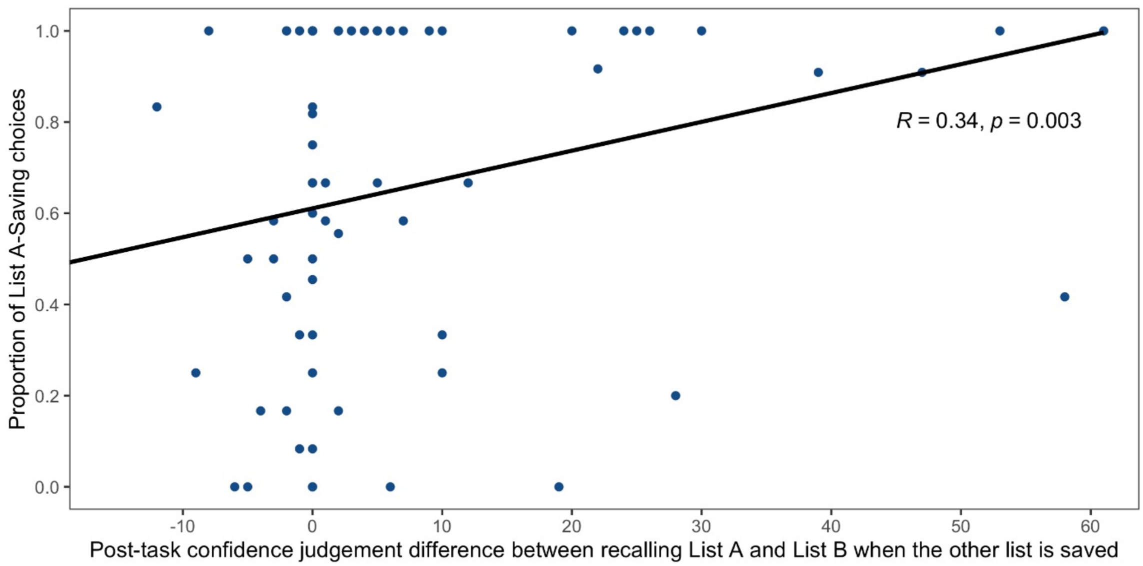

FIGURE 4 Relationship between the proportion of times participants chose to save List A and the difference in their post-task confidence judgments when the other list was saved.

4 | GENERAL DISCUSSION

The main purpose of this study was to conduct a further investigation of the mechanisms underlying the so-called *saving-enhanced memory effect*, whereby offloading the first of two lists of words improves recall not only for that list, but also for the subsequent list (Runge et al., 2019; Storm & Stone, 2015). In particular, Experiment 1 investigated (A) if the saving-enhanced memory effect holds when the offloaded information is presented after the to-be-remembered information and (B) if the saving-enhanced memory effect is affected by the order in which the materials are recalled. Experiment 2 evaluated (A) whether, when given free choice, people prefer offloading rather than relying on internal memory and (B) whether they choose to offload the information that leads to a greater saving-enhanced memory effect. The study was motivated by the concept that a better understanding of the factors that contribute to the saving-enhanced memory effect, as well as those that align offloading strategies with this effect, is important for individuals to derive maximum benefit from offloading strategies in their daily lives.

In Experiment 1, the saving-enhanced memory effect was found for List B when List A was saved and tested second, as in the design used by Storm and Stone (2015), but it was not sustained when List A was tested first. Furthermore, offloading List B did not enhance recall for List A. Previous research has suggested that the saving-enhanced memory effect aligns with the phenomena of LMDF (Runge et al., 2020, 2021; Storm & Stone, 2015), such as the *enhanced encoding* explanation (Pastötter & Bäuml, 2010) and *the reduced interference at recall* hypothesis (Bjork & Bjork, 1996). According to the first account, saving information frees cognitive resources, which would otherwise be utilized for rehearsing that information. These resources, in turn, could then be allocated to learning new information, thus improving performance. In relation to the current study, saving either List A or List B could potentially free cognitive resources. However, in Experiment 1, it was found that saving List B did not produce a saving-enhanced memory effect. This could be explained by the order in which the two lists were memorized. In both experiments, List A was always encoded before List B. So, when participants were presented with a List A-Saving trial, they were aware that List A would be presented again, allowing them to allocate their cognitive resources into encoding List B. On List B-Saving trials, however, participants only knew that List B was saved after encoding both, List A and List B. So, on these trials, participants perhaps allocated resources to both lists, which is perhaps why a saving-enhanced memory effect was not seen when List B was saved. Future studies should explore whether informing participants of the upcoming offloading of List B before encoding List A would be sufficient to elicit a saving-enhanced memory effect for List A when List B is offloaded.

The second mechanism documented by Runge et al. (2020) is that saving information reduces proactive interference at recall as it can be temporarily forgotten and accessed later, reducing interference for recall of subsequently encoded information (Pastötter et al., 2012). Our findings showed that the saving-enhanced memory effect on List B did not hold when List A was tested first. This finding supports the account of saved information reducing proactive interference 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 could have interfered with List A-recall given that the saving cue was only presented to participants after encoding List B. Taken together, our results suggest that these two accounts, although theoretically different, might not be mutually exclusive in the saving-enhanced memory effect. Research in directed forgetting has also proposed various other processes, such as discontinued rehearsal of the first list (e.g., Bjork, 1970) and context change (e.g., Sahakyan et al., 2013), which need to be evaluated by future research.

The current study offers new insights into the degree to which the cognitive processes underlying offloading are similar to (or distinct from) directed forgetting. It does so by examining two factors that have not been previously explored in the cognitive offloading literature but have been investigated in the LMDF field: which list is saved and the order in which the lists are tested. Although research in LMDF has explored whether forgetting the second list would lead to enhanced recall of the first, research on the saving-enhanced memory effect phenomenon has not. Therefore, the present study investigated the consequences of saving List B instead of List A and found that saving the second list (List B in this case) did not result in a saving-enhanced memory effect for the first list (List A in this case). This result is consistent with the prior findings from the LMDF literature, which have indicated no significant difference in recall rates for the List A-items between participants instructed to remember both lists and those instructed to forget List B (e.g., Racsmány et al., 2019). Furthermore, the current study also investigated whether the order in which the two lists were tested would influence the saving-enhanced memory effect. Once again, although this has been investigated in LMDF research, it has not been studied in the cognitive offloading domain. The present study yielded similar results to previous research in LMDF, where the saving-enhanced memory effect of List B was observed only when List B was tested first (Pastötter et al., 2012), and where no recall benefits were found for List A when List B was forgotten, regardless of testing order (Racsmány et al., 2019).

Regarding memory performance of the saved lists, Experiment 1 also 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 than when it was not. Furthermore, this finding also aligns with previous research showing that accuracy is better when participants offload information (see Gilbert et al., 2022, for a review). Seeing as a saved list was tested immediately after restudying it while an unsaved list was tested after a distraction task (or after studying another list plus a distraction task), the retention interval on the Saving trials was much shorter than on the No-Saving trials. So, although the saved list was not presented to participants during the test phase, it was available for restudy right before the test (without any delay). This can be considered as offloading as there was no delay between restudy and test. Furthermore, participants performed an active action to save the information and then retrieve it, which aligns with the definition of cognitive offloading provided by Risko and Gilbert (2016). One limitation of the experimental paradigm adopted in the study is that it does not allow us to fully disentangle the influence of test order and retention interval. Future studies could address this limitation by manipulating the duration of the distractor task to equalize the retention interval for the list that is not saved in the two testing order conditions.

In Experiment 1, participants had to perform an equal number of trials in the three experimental conditions to explore the effect of the two manipulations discussed above. However, in everyday life, we usually have the freedom to decide whether and when to utilize a cognitive offloading strategy. Accordingly, we conducted a second experiment, where participants were given free choice on which of the three strategies to use. We found that when given a choice between saving information or using unaided memory to remember it, participants chose to offload that information. This is consistent with previous research demonstrating that individuals tend to prefer setting reminders rather than relying on internal memory when completing an intention offloading task (see Gilbert et al., 2022, for a review), and with the results presented above, which indicate that accuracy is greater when utilizing cognitive offloading strategies. The novelty of the experiment was not to investigate if participants preferred offloading over relying on their own memory but rather to examine which information they chose to offload when presented with two sets of material to memorize (List A or List B). We found that participants preferred choosing to save List A when it was tested second but did not have a preference towards choosing a specific list when it was tested first. This result shows that participants had a preference for the saving strategy that proved to be the most advantageous in Experiment 1.

The second aim of Experiment 2 was to investigate the role of participants' metacognitive beliefs in saving decisions. Previous literature has shown that offloading behavior is influenced by confidence in our memory abilities (see e.g., Boldt & Gilbert, 2019; Gilbert, 2015; Hu et al., 2019). Although Experiment 2 did not find evidence for the influence of pre-task confidence judgments on participants' offloading strategies, it did find a significant association between post-task metacognitive judgments and participants' decision to offload. Specifically, the difference between participants' post-task confidence ratings for List B-recall when List A was saved and List A-recall when List B was saved correlated with their decision to offload one list over the other. This suggests that in their post-task confidence ratings, individuals likely had some metacognitive insight into their performance. As metacognitive beliefs were likely revised during the task, post-task confidence judgments were more strongly associated with participants' offloading strategy compared to pre-task confidence judgments (see Boldt et al., 2019). Still, participants showed a preference towards offloading the most advantageous list, suggesting that they were able to evaluate and select the best strategy. Perhaps, participants used some other metacognitive beliefs to inform metacognitive control. It should be noted that the task used in the study is rather artificial, and it might be difficult to translate metacognitive beliefs and confidence into a percentage score (see Higham et al., 2016). Furthermore, each confidence rating was asked on a different computer screen, making it potentially harder for the participants to compare one condition to the other. Further studies should be conducted to investigate these possibilities. Yet, it would be interesting to ask participants to make metacognitive predictions at the item level to obtain a more fine-grained overview of the accuracy of participants' metacognitive beliefs.

Given that the results of this study support the notion that offloading information onto external resources can enhance recall, exploring the role of metacognition in this line of research may help improve our use of external tools to support our memory. Future research should also explore how individuals can improve offloading strategies to gain more benefits. For example, the results of the first experiment showed that the saving-enhanced memory effect on List B was only present when the first list (List A) was saved and recalled second. Experiment 2 found that although List A-saving increased when it was tested second, not everyone chose to always save List A. Furthermore, participants only chose to save a list 63% of the time. The reason for this could be that participants were aware that the list they chose to save was only available for restudy on 50% of the trials. The aim of this manipulation was to make the results comparable to the previous studies—that is, Experiment 1 in the current study and Storm and Stone (2015), where participants were uncertain about whether they would have the chance to save the list they were studying as the save button only appeared once the list was encoded. Another possible factor was that the No-Saving trials were faster than the save trials. According to the "soft constraints hypothesis" (Gray et al., 2006), people prefer strategies that take less time. Perhaps participants did not always choose to save a list because it was faster to complete the task using unaided memory. These speculations could also explain why the effect of pre-task confidence judgments on offloading decisions was not found in Experiment 2. Therefore, the results of the metacognitive judgments might not generalize to other situations in which individuals have unrestricted access to the information they choose to save, or where offloading strategies do not significantly impact task completion time. Future research is needed to address these points.

Future research should also evaluate factors contributing to offloading one list over the other in this paradigm to better ascertain factors related to strategy choice. As seen by the results of this study, participants did not always offload one of the two lists, which would have enhanced memory. Furthermore, participants' pre-task metacognitive judgments did not influence their strategy choice. This highlights the importance of interventions aimed at helping sustain the saving-enhanced memory effect. Given the expansion of technology and the convenience of storing information on digital devices, tools can be developed to facilitate memory. These could involve information management tools that offer software-assisted cognitive offloading, capable of automatically storing and representing information based on its relevance to the current task, or making suggestions about what information should be saved based on upcoming activities (see also Gauselmann et al., 2023).

In conclusion, the present study adds to our understanding of how technology is changing our memory by suggesting boundary conditions for the saving-enhanced memory effect. The first experiment found that this effect only holds when the first list (i.e., List A) is offloaded and recalled second. Furthermore, it provides evidence for the use of appropriate metacognitive control strategies. That is, the second experiment found that when given a choice, participants not only chose to save a list but also saved the first list (i.e., List A) when it was recalled second. So, when given a choice between saving two lists, participants tended to choose the list that was most beneficial in terms of recall.

FUNDING INFORMATION

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

CONFLICT OF INTEREST STATEMENT

The authors declare that they do not have competing interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in OSF at https://osf.io/vb8te/ for Experiment 1, https://osf.io/m9ths/ for Experiment 2.

ORCID

Pei-Chun Tsai https://orcid.org/0000-0002-0415-3501
Chhavi Sachdeva https://orcid.org/0000-0002-0074-4371
Sam J. Gilbert https://orcid.org/0000-0002-3839-7045
Chiara Scarampi https://orcid.org/0000-0003-0152-1441

ENDNOTES

[1] Additional follow-up tests were conducted for each reported analyses correcting for multiple comparisons with Bonferroni correction. The results were consistent with those reported in the main text.

[2] Full statistics can be found in the Supplemental Materials.

REFERENCES

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

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

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

Bjork, E. L., & Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness and Cognition*, *5*(1–2), 176–196. https://doi.org/10.1006/ccog.1996.0011

Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior*, *9*(3), 255–268. https://doi.org/10.1016/S0022-5371(70)80059-7

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

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. https://doi.org/10.1038/s41598-019-40681-9

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

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

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*, 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Gauselmann, P., Runge, Y., Jilek, C., Frings, C., Maus, H., & Tempel, T. (2023). A relief from mental overload in a digitalized world: How context-sensitive user interfaces can enhance cognitive performance. *International Journal of Human-Computer Interaction*, *39*(1), 140–150. https://doi.org/10.1080/10447318.2022.2041882

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

Gilbert, S. J., Bird, A., Carpenter, J., Fleming, S. M., Sachdeva, C., & Tsai, P.-C. (2020). Optimal use of reminders: Metacognition, effort, and cognitive offloading. *Journal of Experimental Psychology: General*, *149*(3), 501–517. https://doi.org/10.1037/xge0000652

Gilbert, S. J., Boldt, A., Sachdeva, C., Scarampi, C., & Tsai, P.-C. (2022). Outsourcing memory to external tools: A review of 'intention offloading'. *Psychonomic Bulletin & Review*, *30*, 60–76. https://doi.org/10.3758/s13423-022-02139-4

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.

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.

Higham, P. A., Zawadzka, K., & Hanczakowski, M. (2016). Internal mapping and its impact on measures of absolute and relative metacognitive accuracy. In J. Dunlosky, S. Uma, & K. Tauber (Eds.), *The Oxford handbook of metamemory* (pp. 39–61). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199336746.013.15

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.

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. (2020). 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

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/10.1016/j.jesp.2013.03.013

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.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation*, *26*, 125–173. https://doi.org/10.1016/s0079-7421(08)60053-5

Pastötter, B., & Bäuml, K.-H. (2010). Amount of postcue encoding predicts amount of directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 54–65. https://doi.org/10.1037/a0017406

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

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

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

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*, *85*, 1633–1644. https://doi.org/10.1007/s00426-020-01341-0

Runge, Y., Frings, C., Tempel, T., & Pastötter, B. (2021). Electrophysiological correlates of saving-enhanced memory: Exploring similarities to list-method directed forgetting. *European Journal of Neuroscience*, *54*(6), 6060–6074. https://doi.org/10.1111/ejn.15368

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. (2003). Can encoding differences explain the benefits of directed forgetting in the list method paradigm? *Journal of Memory and Language*, *48*(1), 195–206. https://doi.org/10.1016/S0749-596X(02)00524-7

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

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1064–1072. https://doi.org/10.1037/0278-7393.28.6.1064

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

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Tsai, P.-C., Sachdeva, C., Gilbert, S. J., & Scarampi, C. (2023). An investigation of the saving-enhanced memory effect: The role of test order and list saving. *Applied Cognitive Psychology*, 1–13. https://doi.org/10.1002/acp.4067