AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Learning, Memory, and Cognition

ISSN: 0278-7393

2023, Vol. 49, No. 4, 590–606
https://doi.org/10.1037/xlm0001191

Benefits From Prospective Memory Offloading Depend on Memory Load and Reminder Type

Phil Peper, Durna Alakbarova, and B. Hunter Ball
Department of Psychology, University of Texas at Arlington

Prospective memory (PM) refers to the ability to remember to complete a task at the appropriate moment in the future. Past research has found reminders can improve PM performance in both laboratory and naturalistic settings, but few projects have examined the circumstances when and what types of reminders are most beneficial. Three experiments in the present study tested the influence of reminders in an event-based PM task under different cognitive loads. An additional study examined how effective reminders of different types were. In Experiments 1 (specific targets) and 2 (nonspecific targets), load was manipulated by having participants respond to a single (low load) or multiple (high load) targets. In Experiment 3, the association between target-action word pairs was manipulated by presenting strongly associated pairs (low load) or weakly associated pairs (high load). Experiment 4 used target-action word pairs and varied the type of reminder. Participants in the reminder conditions had target (Experiments 1 and 2), target and action (Experiment 3) or target and/or action (Experiment 4) reminders listed at the top of the screen throughout the PM task. Across the first 3 experiments, it was found that the benefit of reminders was greater under high load than low load conditions. Experiment 4 found that target-action reminders improved PM, while target-only or action-only reminders did not. Importantly, the improvements in PM from reminders occurred without cost to ongoing task performance. Together these results suggest that reminders can be beneficial for reducing PM failures, particularly under high load, without the potential downside of increased effort expenditure.

Keywords: cognitive load, offloading, prospective memory, reminders

Event-based prospective memory (PM) refers to our ability to remember to complete deferred intentions in response to environmental cues, often while busily engaged in various ongoing activities. For example, encountering a medicine bottle while doing chores may serve as an event-based cue to take a medication. Although a single intention may be relatively easy to manage, oftentimes people have multiple concurrent intentions, various cues to respond to, and/or different actions to perform. One potentially efficacious way to deal with the difficulty of maintaining multiple intentions is to offload demands onto the environment (Risko & Gilbert, 2016). Using reminders, such as smartphone alerts for taking a medication or writing to-do lists for groceries, reduces cognitive load and can mitigate capacity limitations that can be potentially reached as intentions mount. Considering the great number of possible forms reminders can take, an important question remains of whether all reminders are equally effective. Although reminders have been shown to be beneficial in both laboratory settings (e.g., Gilbert, 2015a; Guynn et al., 1998; Vortac et al., 1995) and naturalistic settings (Ihle et al., 2012; Schnitzspahn et al., 2020), reminders can be effortful to set up and increase the risk of forgetting should the reminder fail (Kelly & Risko, 2019). The present study aims to explore the efficacy of different reminders in a variety of event-based PM tasks under varied cognitive load.

In the laboratory, event-based PM is typically examined by instructing participants to fulfill an intention embedded within an ongoing task. For example, participants may be instructed to perform an ongoing syllable judgment task in which they decide whether a word on the screen has two or three syllables as quickly and accurately as possible. Prior to performing the task participants form the intention to respond to a certain target word (e.g., the word "apple") or target category (e.g., any fruit) with a unique key response (e.g., press "7") whenever they are encountered during the ongoing task. That is, an intention consists of an association between a target event (e.g., "apple") and a corresponding action (e.g., press "7"). The primary outcome variable is PM performance, which refers to the proportion of PM targets that receive a successful PM response. A secondary measure often assessed is the speed and accuracy of ongoing task responding on nontarget trials during the PM block. Participants are often slower and/or less accurate when possessing an intention compared with when the same task is performed without an intention (Smith, 2003), which is referred to as ongoing task *cost*. This cost to ongoing task

This article was published Online First October 6, 2022.
Durna Alakbarova https://orcid.org/0000-0002-9938-3851
B. Hunter Ball https://orcid.org/0000-0002-9119-7640
Portions of the data were presented at the Virtual Psychonomics 2020 Annual Meeting and the 2020 Armadillo Conference. All data are currently publicly available: https://osf.io/z59ru/.
Correspondence concerning this article should be addressed to B. Hunter Ball, Department of Psychology, University of Texas at Arlington, 501 Nedderman Drive, Arlington, TX 76109, United States. Email: Hunter.Ball@uta.edu

performance is typically thought to reflect diverting attention away from the ongoing task and toward PM processing (but see Heathcote et al., 2015).

Importantly, not all PM intentions incur costs to ongoing task performance. The Multiprocess Framework of PM states that different processes (i.e., monitoring and spontaneous retrieval) can be used to notice targets (Einstein et al., 2005; Shelton et al., 2019). Monitoring is a top-down process that requires controlled attention, such as checking every trial for a PM target (Ball & Brewer, 2018). Owing to the controlled attention needed, monitoring can lead to additional ongoing task costs (Smith, 2003). Spontaneous retrieval is less attentionally demanding and can be achieved through two mechanisms: noticing-plus-search and reflexive-associations (McDaniel & Einstein, 2007). Noticing-plus-search is driven by previous exposure to a target (i.e., intention encoding). Familiarity from previous exposure can increase how fluently a target is processed and provide a discrepancy signal for recognizing the significance of a target. This discrepancy triggers a controlled memory search to identify its source and leads to retrieval of the intended action. With the reflexive-association process, an action that is strongly bound to the target event in memory can be reflexively (automatically) retrieved upon fully processing the relevant features of the target (Einstein & McDaniel, 2005). PM performance can thus be improved by encouraging more effective or consistent monitoring, enhancing the recognition signal of a target, or strengthening the association between the PM target and action. Importantly, the latter two mechanisms should not incur additional ongoing task cost.

Alternatively, Marsh et al. (2003) outlined a microstructure of PM target detection that consists of four subprocesses, including (a) *recognition* of a target as related to the PM task, (b) *verification* that the PM intention included the target and current context, (c) *retrieval* of the action associated with the target, and (d) *coordination* of the PM response with the ongoing task response. The microstructure and the Multiprocess Framework have similar assumptions. First, both monitoring and the discrepancy signal influence how well a target is recognized (Breneiser & McDaniel, 2006; Smith, 2003). Second, the search for the source of the discrepancy signal is related to verification of the target, such that the target must be verified as relevant to the PM intention on being recognized (Marsh et al., 2003). Last, the strength of the association between the target and action affects how easily the action is retrieved on verifying the target (Cook et al., 2014; McDaniel et al., 2004).

The specificity[1] of the targets can lead to different processes being used to notice the PM targets. Specific targets are encoded explicitly (e.g., *apple*) as they appear in the ongoing task and are often detected at high rates with minimal cost to ongoing task performance, suggesting they are retrieved spontaneously (Scullin et al., 2010). In contrast, nonspecific targets (e.g., exemplars from the category *fruits*) necessitate monitoring to actively assess stimuli for PM-relevant features that were not explicitly encoded, resulting in higher costs and lower target detection (Einstein et al., 2005; Marsh et al., 2003). Because PM monitoring places demands on working memory (Brewer et al., 2010), this suggests that nonspecific targets produce greater cognitive load than specific targets.

Research has found that manipulations that increase the memory load can reduce PM performance. For example, Marsh et al. (2003) found that increasing targets from four to eight reduced PM performance, whereas Cohen et al. (2008) found that increasing the number of targets increased the ongoing task cost. Another way to increase the memory load of a PM task is assigning each target a unique action and manipulating the association between the target and the target action. For example, participants may learn a series of paired associates (e.g., *salt* – *pepper*) with the intention that, any time they see the target (e.g., *salt*) during the ongoing task, they should type out the action word that was paired with it (e.g., *pepper*). Several studies have found that manipulating the degree of association between the target-action pairs impacts PM performance (e.g., Cook et al., 2014; McDaniel et al., 2004). For example, PM performance is much higher with strongly associated pairs (e.g., *salt* – *pepper*) than weakly associated pairs (e.g., *candle* – *elbow*). Better PM performance for strong associations occurs because participants need less controlled retrieval to recall the associated pair on noticing the target (Einstein & McDaniel, 2010). In contrast, retrieval of a weak pair requires a controlled search process. Thus, like increasing the number of targets, weakening the association between PM targets and actions can increase memory demands.

One way to reduce the cognitive load of a task is to offload goal-related contents onto the environment (Risko & Gilbert, 2016). For example, people can create to-do lists, update online calendars with future appointments, or set a gym bag next to their front door to not forget it before leaving for work. Einstein and McDaniel (1990) conducted the first experiment on PM and reminders in the laboratory using the paradigm like that described above. They found that when participants wrote down the targets on a piece of paper, PM performance was higher than a condition with no reminders. Vortac et al. (1995) compared the effectiveness of reminders available at different time periods during an air traffic control PM task. Specifically, they compared reminders available before the PM target occurred (but not during), throughout the PM task, and selectively available in a small window around the PM trial. They found that PM target reminders were effective as long as they were available during the PM trial (but see Loft et al., 2011, for evidence of habituation to always-available reminders). Consistent with Vortac et al. (1995), Guynn et al. (1998) found that brief reminders of the target alone or action alone presented approximately one-minute before the PM trial provided no benefit to PM performance. However, reminders of the target *and* action combined improved PM performance. Chen et al. (2017) manipulated ongoing task load by having participants do a 1-back or a 2-back task and found that target reminders improved PM performance regardless of task load. Interestingly, they found reminders improved ongoing task performance under high ongoing task load. Finally, Henry et al. (2012) used a Virtual-Week PM task and examined PM performance with self-generated reminders, experimenter-generated reminders, and without reminders. They found that reminders improved PM but observed no difference between self- and experimenter-generated reminders, suggesting reminders are helpful regardless of their source. Importantly, however, although

[1] Specificity is similar to the idea of target "focality," but it makes no assumptions about the overlap in processing between the PM targets and ongoing task. Conceptually, however, specific and focal target detection may rely on similar processes, whereas nonspecific and nonfocal target detection may rely on similar processes.

these studies have been critical in demonstrating when reminders are effective, none of these studies manipulated the memory load of the PM task. By manipulating memory load, the current study can examine the scenarios in which reminders are most effective (e.g., when there are many intentions to remember) and can further pinpoint the mechanism by which reminders improve PM.

The only studies examining the effectiveness of reminders under varied PM load in an adult sample used the *intention offloading task*. In the task, participants are instructed to drag centrally located circles to the bottom of the screen (Gilbert, 2015a; Gilbert et al., 2020). Each circle is either numbered or alphabetized, and the circles are dragged sequentially. A random circle is periodically flashed a certain color (e.g., blue). When the circle is reached in the alphabetical sequence, it must be dragged to a different side of the screen (e.g., left) that matches the flashed color (e.g., blue). Participants can set reminders by dragging the target circle next to its destination as soon as it flashed, reminding participants of the unique response required. Otherwise, the intention must be remembered internally (i.e., without setting a reminder). The primary finding in this paradigm is that performance is considerably higher when participants set reminders (e.g., Cherkaoui & Gilbert, 2017; Gilbert, 2015a, 2015b; Gilbert et al., 2020). Critically, reminders are used more often and show a greater benefit when the number of target circles within a trial block is high (i.e., with increased cognitive load; see Bulley et al., 2020; Redshaw et al., 2018; for similar results in children).

Although the findings from the intention offloading task are important, the delay intervals during the intention offloading task occur over the course of 10 to 20 seconds, which is much shorter than typical PM tasks. This means that the intention can be maintained in working memory until execution is appropriate (i.e., a vigilance task; Graf & Uttl, 2001). Prior studies have also not calculated ongoing task costs (i.e., the time it takes to drag nontarget trials), which is a central to many theories of PM. Given these concerns, the theoretical processes and the microstructure of PM derived from the typical lab-based PM paradigm (Einstein & McDaniel, 1990) have largely been unexplored in the intention offloading task. It is therefore important to extend the findings of Gilbert and colleagues to a more typical event-based PM task that allows for a more targeted theoretical investigation of how reminders can improve PM.

The Present Study

The present study sought to better understand the mechanisms by which reminders may benefit prospective remembering. A traditional PM paradigm was used in which PM targets were embedded in an ongoing syllable judgment task. The type of PM targets and reminders was manipulated across experiments. PM targets included specific targets requiring a single action (Experiment 1), nonspecific targets requiring a single action (Experiment 2), and specific targets each requiring different actions (Experiments 3 and 4). Reminders involved displaying PM target (Experiments 1 and 2) or PM target and/or action (Experiments 3 and 4) information at the top of the computer screen throughout the duration of the PM task (see Figure 1). By providing reminders for participants, this reduced potential metacognitive biases associated with suboptimal reminder setting that may otherwise influence performance (e.g., Gilbert et al., 2020). Memory load (e.g., one vs. four PM targets) was manipulated in each experiment to assess the efficacy of reminders under different demands. The general hypothesis was that reminders would be more effective under high load. Importantly, however, this could occur for different reasons.

The first way by which reminders may improve PM is by enhancing target *recognition* through greater monitoring or spontaneous retrieval. Reminder checking may cause participants to monitor more thoroughly or consistently for targets or increase the fluency of the targets (in contrast to the novel ongoing task stimuli) and strengthen the discrepancy signal. Alternatively, reminders may be used to offload *verification* process. On recognizing the target, one could look to the reminder to verify that the target is relevant to the intention as an alternative to controlled memory search. Finally, reminders may make it easier to *retrieve* the intention by strengthening the association between the target and action. It should be noted that these processes are not necessarily mutually exclusive. However, in each case reminders should be more beneficial under high demands, because greater memory demands (e.g., multiple targets, low target-associations) make detecting the targets more difficult and forgetting the intended action should be more likely.

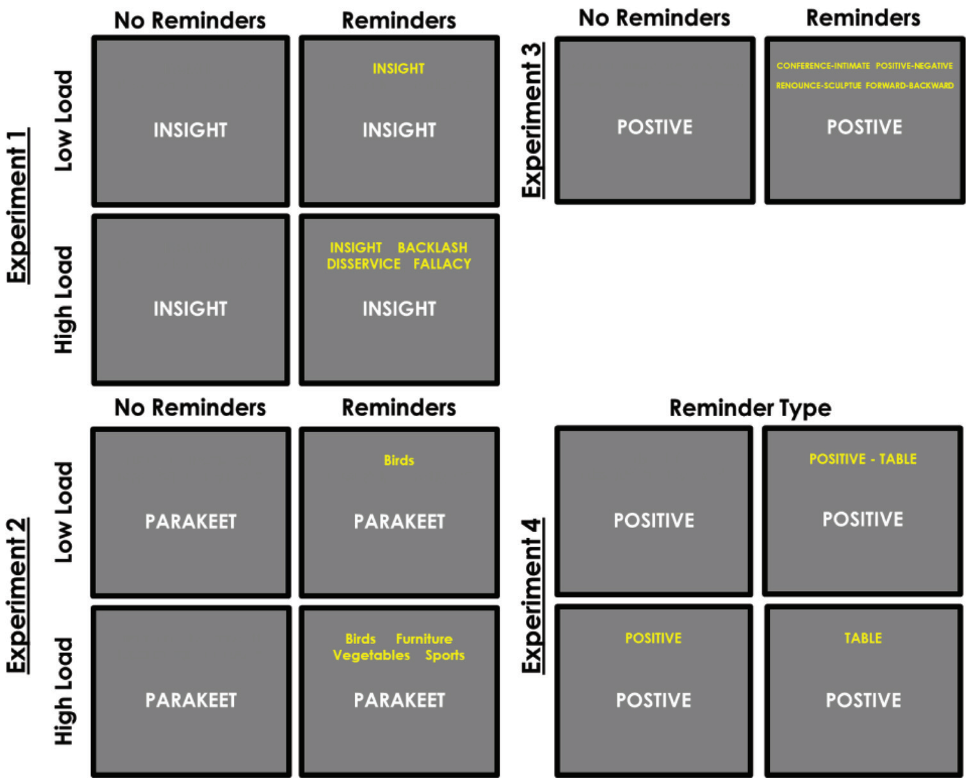
All research reported herein was conducted using appropriate ethical guidelines and was approved by the Institutional Review Board at the University of Texas at Arlington. We report how we determined our sample size, all data exclusions, and all manipulations.[2] Links to data files and each preregistration are listed in the introductions to each respective experiment.

Experiment 1

Experiment 1 examined the influence of specific target reminders on PM performance. Participants responded to one (low load) or four (high load) specific PM targets, with or without the use of reminders. In the reminder conditions, the single target (low load) or all four targets (high load) were presented at the top of the screen (see Figure 1). In the no reminder conditions, participants had to rely on their own memory to notice the targets. Previous research has shown that a single specific prospective memory target can be

[2] It should be noted that we initially conducted three other experiments (a, b, and c) that are not included in the current article (called "PMReminders Initial Registration" at https://osf.io/z59ru/). In these studies, load was manipulated by comparing performance for four (low load) versus eight (high load) specific targets for the first two (collected through SONA and then MTurk) and five (low load) versus ten (high load) specific targets for the third (collected on MTurk). The MTurk studies were conducted to ensure similar patterns of results were found in person and online but were not preregistered owing to an oversight on the first author's part while research was shifted online amid the early stages of COVID-19–related precautions. Although reminders improved PM performance, there was not the anticipated interaction between load and reminders (these data are available on request). The experiments were deemed flawed for the purpose of the study after we realized that four targets were actually *high* load, thereby reducing the validity of the load manipulation. Experiment 2 of the current study (nonspecific targets) changed the load manipulation (i.e., 1 vs. 4) and found the anticipated interaction. Experiment 1 of the current study (specific targets) was then re-ran using the same load manipulation (i.e., 1 vs. 4), followed by Experiment 3 (i.e., target-action pairs). Because of this ordering, there is some confusing text in the OSF documents, as Experiment 1 of the current study states that it is "Experiment 3" in the text and Experiment 3 of the current study states that it is "Experiment 4." The filenames (e.g., "Experiment 1 Preregistration") have been relabeled to directly map onto how they are reported in the current manuscript (e.g., "Experiment 1").

Figure 1
Task Appearance in Each Condition for All Four Experiments



Note. Squares on the left and right for each experiment represent the appearance of the prospective memory (PM) task for the no reminder and reminder conditions, respectively. In Experiments 1 and 2, the bottom squares represent task appearance for the high load conditions. Load was manipulated within subjects in Experiment 3. The type of reminder was manipulated in Experiment 4 and were thus combined in the figure: target-action reminders in the top-right, target reminders bottom-left, and action reminders bottom-right. See the online article for the color version of this figure.

detected at relatively high rates with minimal costs to ongoing task performance. Using multiple specific targets places greater demands on working memory maintenance as evidenced by increased ongoing task costs (Cohen et al., 2008), but retrieval processes presumably still operate similarly with a single action associated with all targets (Humphreys et al., 2020; Strickland et al., 2022).

Given that retrieval processes operate similarly when only adding to the target load, increasing the number of targets should increase the demand on target recognition (Wesslein et al., 2014). Following encoding of a specific word (e.g., *parakeet*), reexperiencing this item via reminder checking during the retrieval phase may strengthen a discrepancy signal due to increasing processing fluency, which automatically stimulates search for the source of the discrepancy. According to this view, increased discrepancy should be particularly beneficial when there are multiple PM targets, because a single specific target already produces a strong discrepancy signal. Reminders that increase discrepancy should not incur additional costs to ongoing task performance. Alternatively, with a reminder constantly available, participants may be more likely to focus attention on task-relevant goals (i.e., have fewer lapses of intention; West & Craik, 1999) and monitor more consistently or thoroughly throughout the task. According to this view, reminders should be associated with worse ongoing task processing indicative of more attention devoted to the PM task, or greater perceived subjective importance of the intention on a postexperimental questionnaire. Experiment 1 preregistration information can be found at https://osf.io/9tx4a and data can be found at https://osf.io/fvk9g (Peper et al., 2022b).[3]

Method

Participants

An a priori power analysis based on a medium effect size (η_p^2 = .06) recommended a sample size of 175 for an α of .05 and power of .80. We set the final sample size to 200 with anticipation that some participants will be removed for meeting exclusionary criteria. Participants (age 18–37 years; M = 19.12, SD = 2.45) were undergraduates at the University of Texas at Arlington recruited through SONA systems for class credit. Of the total 200 participants collected, 11 were excluded for meeting exclusionary criteria

[3] As described in footnote 2, the text of the preregistration document says "Experiment 3" although it corresponds to Experiment 1 in the current article.

detailed below, for a final sample size of 189. Participants were randomly assigned to one of four conditions, including the no reminder low load (N = 47), no reminder high load (N = 50), reminder low load (N = 43), and reminder high load (N = 49).

Design

A 2 (load: low vs. high) × 2 (reminder: reminder vs. no reminder) between-subjects design was used, such that participants either learned one (low load) or four (high load) PM targets with (reminder condition) or without (no reminder condition) the use of reminders.

Materials

Ongoing task stimuli were selected from the English Lexicon Project (Balota et al., 2007). These consisted of 252 words that were seven to 10 letters in length, half of which had two syllables and half of which had three syllables. There was an equal number of two- and three-syllable words of each letter count. An additional four words were selected as PM targets. Two of the PM targets were two-syllable words (*insight* and *backlash*) and two were three-syllable words (*fallacy* and *disservice*). The stimuli were presented in uppercase white font at the center of the screen on a dark gray background. Participants in the reminder conditions had the target(s) listed at the top of the screen in yellow font.

All data collection was completed online. Experimental procedures were developed on and presented with PsychoPy3 software. QuestionPro survey software obtained participant consent and randomly assigned participants to each condition. The PsychoPy3 experiment was hosted on Pavlovia.com.

Procedure

The experiment involved participants performing a syllable judgment ongoing task with PM targets embedded. Participants completed demographics questions, practiced the ongoing task, received the intention instructions (and were quizzed), learned a set of targets (and were quizzed), performed the actual task, and completed a postexperimental questionnaire. For the ongoing syllable judgment task, participants were required to make two versus three syllable judgements about English words presented on a computer. Participants pressed the "F" key for two-syllable words (e.g., *boredom*) and the "J" key for three-syllable words (e.g., *cynical*). After each judgment, a brief (500ms) fixation cross appeared before another word stimulus was presented.

Practice Block Phase. After reading instructions for the ongoing task, participants completed a 20-trial practice block and received accuracy feedback after each trial. Participants were only allowed to proceed after achieving 75% accuracy or greater on the practice. Afterward, participants performed another practice block (40 trials) without feedback.

Attention Check Phase. Participants provided gender information after the practice block phase along with a filler question. There was then an attention check that instructed the participant to "Press the 2-key to indicate you are paying attention." Attention check failures resulted in immediate exiting of the experiment and participants not receiving credit.

Intention Instruction Phase. On completing the practice phase, participants received instructions for the upcoming PM task. Participants were instructed that they were going to learn a single word or list of words that were to later appear during the syllable judgment task. The PM intention was to make a special response (press the "7" key) whenever they encountered the PM target(s). They were to press the "7" key *instead* of making their ongoing task response. However, participants were told that the primary objective was still to perform the ongoing task as quickly but as accurately as possible. A brief instructions quiz was then presented to each participant with two questions. The first question asked about the goal of the intention (i.e., look for specific words). The second question asked for the PM response (i.e., press the "7" key). Participants had to get every question correct before proceeding. If they answered a question wrong, they had to reread the instructions to ensure proper encoding of the PM task.

Target Learning Phase. Participants then studied one (low load condition) or four (high load condition) PM targets for five seconds each. Their memory for the targets was then quizzed using a two-alternative forced choice test. A PM target and a control-matched word (i.e., matched in letter and syllable length) were presented on the left and right side of the screen. Half of the time the correct answer appeared on the left side of the screen and half of the time it appeared on the right. Participants pressed either the left or right arrow key to indicate which word they studied previously. Participants had to achieve 100% accuracy before proceeding. If there was an incorrect answer, at the end of the recognition quiz they were required to restudy all the targets and take the quiz again.

Reminder Phase. After achieving 100% accuracy, the PM instructions were reiterated for all participants. The no reminder conditions then began the distractor task. The reminder condition was additionally instructed, "To help you remember the studied words, they will be listed in the upper part of your screen in yellow during the next syllable judgment task." The reminder conditions then began the distractor task.

Distractor Phase. Participants completed arithmetic problems involving multidigit addition and subtraction for two-minutes before the PM block.

PM Block Phase. Before beginning the PM block, participants received instructions that reiterated only the ongoing task instructions. The ongoing task consisted of 200 trials with word type (two vs. three syllable) randomly presented. Every 25 trials a PM target was presented. In the low load conditions, one of the four PM targets was randomly selected and then presented eight times. In the high load conditions, the four cues were randomly presented two times each, once in the first half and once in the second half of the PM block.

For the no reminder conditions, the task appearance mirrored the second ongoing task practice block. In contrast, for the reminder conditions, participants additionally had the single learned PM target (low load) or all four PM targets (high load) listed in yellow font at the top of the screen (see Figure 1). The target reminders remained on the screen for the entire duration of the PM block.

Postexperimental Questionnaire Phase. After the PM block, participants completed postexperimental questions assessing their retrospective memory for the PM task. They were first asked to freely recall the intention. They were then given a multiple-choice test to identify the PM intention (i.e., "look for specific words") and the response (i.e., "press 7"). Participants then answered two 5-point Likert scale questions, one on the perceived importance of the ongoing task and the other on the perceived importance of the PM task (1 = *Not at all important*, 5 = *Very important*). An attention check followed the importance questions instructing the participant to "Press the 3-key to indicate you are paying attention." The reminder conditions were then asked a 5-point Likert scale question about how frequently they checked the reminder (1 = *Not at all*, 5 = *All the time*). Recognition memory for the PM targets was assessed using the same two-alternative forced choice test as in the target learning phase. Last, participants were asked whether they wrote down the PM targets. Participants were informed that reporting cheating would not influence their credit compensation.

Exclusionary Criteria

Exclusionary criteria for the most part were preregistered. Deviations are specified and explained below. The preregistration specified that participants who failed an attention check were excluded from analyses. However, attention check failures resulted in immediate termination of the study, so these participants were not included in the exclusions below. Participants who failed the attention checks had incomplete data and were not counted toward the sample total. Participants were excluded for the following: (a) zero scores on PM performance *and* failing to recognize the PM response key on the postexperimental questionnaire (i.e., retrospective memory error; $n = 1$); (b) reporting cheating (writing down PM targets; $n = 4$); (c) performing below 50% on the math distractor ($n = 1$); (d) achieving less than 50% ongoing task accuracy for either two- or three-syllable words in both the control and PM block ($n = 1$); (e) duplicates ($n = 4$); (f) false alarms ($n = 0$). Participants meeting first exclusionary criterion (zero PM performance *and* forgetting the PM task instructions) represented a failure in retrospective memory rather than an PM error. Although not preregistered, we set an additional exclusionary criterion of false alarms on greater than 15% of trials (greater than 30 false alarms). Participants excluded for meeting this criterion were rare (only two in Experiment 4), and their removal did not fundamentally alter any results. The false alarm criterion was applied to all experiments for uniformity.

Results

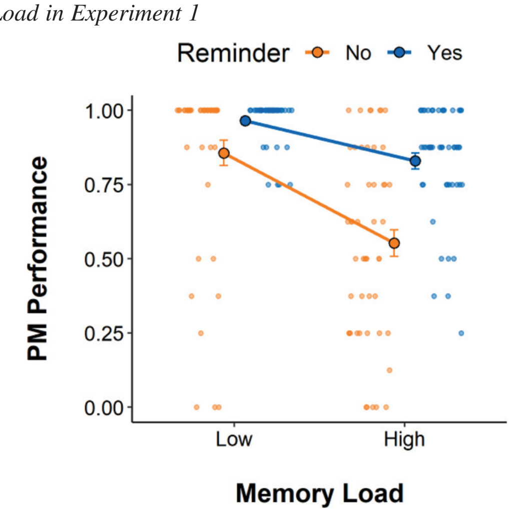
Separate 2 (load: low vs. high) × 2 (reminder: reminder vs. no reminder) between-subjects ANOVAs were conducted for PM performance, ongoing task performance, and recognition memory. For probing interaction effects of reminder by load, we used Bonferroni-corrected p values set to .025 for the two comparisons between the reminder condition and the no-reminder control under high load and low load. A successful PM response involved pressing the "7" on target trials in lieu of making ongoing task response. Late responses were calculated as a proportion of PM responses on trials immediately after a target. These were infrequent (overall $M = .001$, $SE < .001$) and not counted as correct (e.g., Marsh et al., 2002). PM performance was calculated as a proportion of successful PM responses of eight PM target trials. False alarms were calculated as the total number of PM responses made on nontarget trials, excluding late responses. Ongoing task performance during the PM block included accuracy and mean response times. The first three ongoing task trials and the three ongoing task trials following a PM trial were excluded from analyses (Ball & Bugg, 2018). The three ongoing task trials after a PM trial were removed to avoid aftereffects of larger response times following PM targets (Meier & Rey-Mermet, 2012). Ongoing task reaction times were calculated for correct responses only. Trials falling ±3.0 standard deviations from an individual's mean response time were excluded from analyses. Recognition memory refers to the proportion of correctly identified PM targets on the recognition test at the end of the experiment. The aforementioned data analyses were preregistered, except for the Bonferroni corrections. Bonferroni corrections were preregistered for Experiment 4 and applied to follow-up comparisons in all experiments for uniformity. Exploratory analyses were also preregistered and conducted for perceived PM task importance. While not preregistered, exploratory analyses were conducted for frequency of reminder checking.

PM Performance

Proportion Correct. The analysis for PM performance (Figure 2; Table 1) revealed that PM was higher with reminders, *Reminder:* $F(1, 185) = 29.99$, $p < .001$, $\eta_p^2 = .139$, and with low cognitive load, *Load:* $F(1, 185) = 39.09$, $p < .001$, $\eta_p^2 = .174$. There was also an interaction between the two, *Reminder × Load:* $F(1, 185) = 5.69$, $p = .015$, $\eta_p^2 = .031$. This interaction reflects that while reminders improved PM in both the low load, $F(1, 88) = 5.49$, $p = .021$, $\eta_p^2 = .059$, and high load, $F(1, 97) = 28.15$, $p < .001$, $\eta_p^2 = .225$, conditions, the improvements from reminders was greater under high load.

False Alarms. The analysis of false alarms on nontarget trials (see Table 1) revealed that reminders reduced false alarm rates, *Reminder:* $F(1, 185) = 10.00$, $p = .002$, $\eta_p^2 = .051$. There was no effect of cognitive load, *Load:* $F(1, 185) = 1.20$, $p = .275$, $\eta_p^2 = .006$, and there was no interaction between the two, *Reminder × Load:* $F < 1$.

Figure 2
Prospective Memory Performance as a Function of Reminder and Load in Experiment 1



Note. Prospective memory (PM) performance refers to the proportion of PM targets detected (out of eight). Large circles indicate mean performance, whereas the smaller circles reflect individual data points to illustrate the distribution of scores in each condition. Error bars reflect standard error. See the online article for the color version of this figure.

Table 1
Means and Standard Errors for Prospective Memory Performance, Ongoing Task Performance, and Recognition Memory

	Prospective memory		Ongoing task performance		Recognition memory
Experiment and condition	Proportion correct	False alarms	Accuracy	RT (ms)	Target hits
Experiment 1					
Low - No reminder	0.86 (0.04)	1.51 (0.62)	0.94 (0.01)	1,688 (100)	1.00 (0.00)
High - No reminder	0.55 (0.04)	0.88 (0.22)	0.93 (0.01)	1,551 (64)	0.99 (0.01)
Low - Reminder	0.97 (0.01)	0.19 (0.08)	0.94 (0.01)	1,508 (81)	1.00 (0.00)
High - Reminder	0.83 (0.03)	0.08 (0.06)	0.94 (0.01)	1,747 (97)	1.00 (0.00)
Experiment 2					
Low - No reminder	0.51 (0.06)	0.97 (0.35)	0.94 (0.01)	1,489 (77)	1.00 (0.00)
High - No reminder	0.11 (0.03)	1.09 (0.39)	0.95 (0.01)	1,621 (72)	1.00 (0.00)
Low - Reminder	0.58 (0.06)	1.39 (0.46)	0.93 (0.01)	1,553 (99)	1.00 (0.00)
High - Reminder	0.47 (0.06)	2.21 (0.47)	0.92 (0.01)	1,845 (94)	0.99 (0.01)
Experiment 3					
Low - No reminder	0.59 (0.06)	0.42 (0.15)	0.94 (0.01)	1,685 (58)	0.99 (0.01)
High - No reminder	0.27 (0.06)	—	—	—	—
Low - Reminder	0.80 (0.05)	0.02 (0.02)	0.95 (0.01)	1,537 (66)	0.99 (0.01)
High - Reminder	0.69 (0.05)	—	—	—	—
Experiment 4					
No reminder	0.30 (0.05)	0.29 (0.11)	0.94 (0.01)	1,675 (79)	0.99 (0.01)
Target-action reminder	0.67 (0.04)	0.00 (0.00)	0.93 (0.01)	1,630 (76)	1.00 (0.00)
Target reminder	0.24 (0.05)	0.41 (0.39)	0.93 (0.01)	1,551 (92)	0.98 (0.01)
Action reminder	0.46 (0.05)	0.79 (0.17)	0.93 (0.01)	1,661 (68)	1.00 (0.00)

Note. Standard errors in parentheses; prospective memory proportion correct are in proportions (of 1) and false alarms are in frequency; false alarms, ongoing task performance, and recognition memory for Experiment 3 refer to the no reminder and reminder conditions overall, as load was manipulated within-subjects.

Ongoing Task Performance

Accuracy. For ongoing task accuracy (see Table 1), there was no effect of reminder, *Reminder:* $F < 1$, or load, *Load:* $F < 1$, and there was no interaction between the two, *Reminder* × *Load:* $F(1, 185) = 1.38$, $p = .242$, $\eta_p^2 = .007$.

Response Time. For response times (see Table 1), there was no effect of reminder ($F < 1$) or load ($F < 1$), but there was an interaction between the two, *Reminder* × *Load:* $F(1, 185) = 4.68$, $p = .032$, $\eta_p^2 = .025$. This interaction reflects that response times were nominally, but not significantly, faster with reminders than without reminders under low load, $F(1, 88) = 1.89$, $p = .172$, $\eta_p^2 = .021$, and response times were nominally, but not significantly, slower with reminders than without reminders under high load, $F(1, 97) = 2.88$, $p = .093$, $\eta_p^2 = .029$.

Recognition Memory. The analysis of recognition memory (see Table 1) found no significant effect of reminder, *Reminder:* $F(1, 185) = 2.78$, $p = .097$, $\eta_p^2 = .015$, no effect of load, *Load:* $F(1, 185) = 2.78$, $p = .097$, $\eta_p^2 = .015$, and there was no interaction between the two, *Reminder* × *Load:* $F(1, 185) = 2.78$, $p = .097$, $\eta_p^2 = .015$.

Exploratory Analyses

PM Importance. For perceived importance of the PM task (see Table 2), there was no effect of reminder (*Reminder:* $F < 1$) or load (*Load:* $F < 1$) on PM task importance, and there was no interaction between the two (*Reminder* × *Load:* $F < 1$).

Reminder Checking. The analysis of reminder checking frequency was conducted only on the two reminder conditions (see Table 2). Participants reported checking the reminder more frequently in the high load condition than in the low load condition, $F(1, 90) = 43.59$, $p < .001$, $\eta_p^2 = .326$.

Discussion

Experiment 1 tested the effect of reminders on PM as a function of cognitive load using specific targets. As predicted, reminders improved PM performance, particularly under high load. This is consistent with previous research showing that reminders provide a greater benefit when the number of targets is high compared with a single target (Gilbert, 2015a). Of note, the benefit to PM

Table 2
Means and Standard Errors for Prospective Memory Importance and Reminder Checking Frequency

Experiment and condition	Prospective memory importance	Reminder checking frequency
Experiment 1		
Low - No reminder	4.23 (0.12)	—
High - No reminder	3.98 (0.15)	—
Low - Reminder	4.16 (0.14)	1.54 (0.11)
High - Reminder	4.16 (0.12)	2.55 (0.11)
Experiment 2		
Low - No reminder	3.87 (0.17)	—
High - No reminder	3.20 (0.20)	—
Low - Reminder	3.76 (0.20)	1.92 (0.16)
High - Reminder	3.79 (0.16)	3.18 (0.17)
Experiment 3		
No reminder	3.70 (0.15)	—
Reminder	3.77 (0.14)	3.07 (0.17)
Experiment 4		
No reminder	3.83 (0.16)	—
Target-action reminder	3.71 (0.14)	3.36 (0.14)
Target reminder	3.71 (0.14)	3.24 (0.17)
Action reminder	3.84 (0.16)	3.10 (0.13)

Note. Standard errors in parentheses; no reminder checking data collected for no reminder conditions.

performance did not occur due to changes in recognition memory or increased perceived importance of the PM task. However, it is possible that the greater reminder benefit under high load could be partially driven by the increase in reminder checking frequency. Importantly, the benefit of reminders came at no cost to ongoing task performance. There was a significant load by reminder interaction on response times, but reminders did not significantly influence response times within each load condition. The fact that reminders improved PM without significantly impacting ongoing task performance suggests that specific target reminders did not impact monitoring. Specific reminders therefore appear to improve target recognition by increasing the discrepancy signal of the targets. Whereas Cohen et al. (2008) found increasing the number of targets slowed ongoing task responding, Wesslein et al. (2014) did not, so the lack of load effects on ongoing task performance was not altogether surprising.

It should be noted that performance was very high in the low load condition, consistent with previous research (Einstein et al., 2005; Hicks et al., 2017; Marsh et al., 2003). It is therefore possible that the benefits with reminders in the low load condition were limited by ceiling effects. Experiment 2 was designed as a conceptual replication using nonspecific targets that should remove ceiling effects. Additionally, nonspecific reminders cannot increase the fluency (and therefore the discrepancy) of a target (i.e., recognition). Showing benefits from nonspecific reminders may provide insights into other mechanisms by which PM can be improved.

Experiment 2

The purpose of Experiment 2 was to conceptually replicate Experiment 1 using nonspecific (i.e., categorical) PM targets, which incur ongoing task costs due to monitoring and produce lower PM performance (Smith, 2003; Smith & Bayen, 2004, 2005). Load was manipulated in a similar manner as Experiment 1 with one (low load) or four (high load) nonspecific targets (i.e., category words). Those in the reminder conditions had the nonspecific target(s) listed at the top of their screen (see Figure 1). As with Experiment 1, we predicted that reminders would improve PM performance, particularly under high cognitive load.

Nonspecific targets are not encoded explicitly and therefore do not produce a familiarity-induced discrepancy signal. Likewise, nonspecific *reminders* of a general category or categories (e.g., *birds*) do not include the features of the target as it appears in the ongoing task (e.g., *parakeet*), meaning that they cannot improve target recognition by enhancing discrepancy. However, they could still influence recognition by increasing monitoring, which may be particularly beneficial under high load. According to this view, improvements to PM performance from reminders should be associated with increased ongoing task cost and/or greater perceived importance of the PM intention. Alternatively, participants could use nonspecific reminders to verify the relevance of the target for the PM task by checking the reminder rather than verifying relevance with internal processes. According to this view, reminders should benefit PM performance without any additional influence on ongoing task performance. Experiment 2 preregistration information can be found at https://osf.io/7cxvs and data can be found at https://osf.io/jrbaf (Peper et al., 2022b).

Method

Participants

An a priori power analysis based on the effect size from pilot data recommended 120 participants across the four conditions. Owing to experimenter error, 160 participants were collected. All participants (age 18–40 years; $M = 19.64$, $SD = 3.53$) were undergraduate students at the University of Texas at Arlington that received class credit for participation. Of the 160 participants collected, 12 participants were excluded for meeting exclusionary criteria detailed below, leaving the final sample size at 148. Participants were randomly assigned to one of four conditions, including the no reminder low load ($N = 37$), no reminder high load ($N = 35$), reminder low load ($N = 38$), and reminder high load ($N = 38$).

Design and Materials

A 2 (load: low vs. high) × 2 (reminder: reminder vs. no reminder) between-subjects design was used, such that participants either learned one (low load) or four (high load) PM targets with (reminder condition) or without (no reminder condition) the use of reminders. Ongoing task stimuli were the same as Experiment 1. Nonspecific category targets were chosen from category norms (Van Overschelde et al., 2004). The four categories were *birds*, *vegetables*, *furniture*, and *sports*. The three most common exemplars from each category were excluded. There was a total of 36 category exemplars. The bird exemplars were *bluejay*, *cardinal*, *flamingo*, *sparrow*, *seagull*, *parakeet*, *penguin*, and *woodpecker*. The vegetable exemplars were *broccoli*, *spinach*, *potato*, *zucchini*, *pepper*, *cucumber*, *cabbage*, and *lettuce*. The furniture exemplars were *loveseat*, *dresser*, *nightstand*, *ottoman*, *recliner*, *armoire*, *bookshelf*, and *cabinet*. The sports exemplars were *baseball*, *running*, *swimming*, *volleyball*, *bowling*, *badminton*, *wrestling*, and *lacrosse*.

QuestionPro survey software was used to obtain consent and randomly assign participants to each condition. The experiment was developed with PsychoPy3 and hosted on Pavlovia.com in the same manner as Experiment 1.

Procedure

The overall procedure in Experiment 2 matched that of Experiment 1, except the PM targets were nonspecific (i.e., categorical) instead of specific words. Participants learned a single (low load) or four (high load) category words. They were instructed to make a PM response (press the "7" key) whenever they saw an exemplar (e.g., *parakeet*) from a learned category (e.g., *birds*) in the subsequent syllable rating task. Eight PM targets appeared in the PM block, with one occurring every 25th trial. In the low load condition, one of the four categories was randomly selected as the PM target for each participant and each of the eight exemplars was randomly presented during the ongoing task once. In the high load conditions, all four of the categories were selected as PM targets. Two exemplars were randomly selected from each category and presented during the PM block. One exemplar from each of the four categories was randomly selected to appear among the first four PM targets. Then the second exemplar from each category appeared as one of the last four PM targets. The category order and the exemplar from each category was randomized, meaning that across all participants in the high load conditions, each of the exemplars had an equal chance of appearing in the PM block. The reminder conditions had the single category (low load) or all four category (high load) names (e.g., *birds*) listed at the top of the screen in yellow font (see Figure 1).

Exclusionary Criteria

Participants were excluded for the following: (a) failing to have any PM hits *and* failing to recall the PM response key during the postexperimental questionnaire (n = 4); (b) reporting cheating (writing down PM targets; n = 1); (c) performing below 50% on the math distractor (n = 0); (d) achieving less than 50% in ongoing task accuracy for either two- or three-syllable words in both the control and PM block (n = 4); (e) duplicates (n = 3); (f) false alarms (n = 0). Neither math distractor performance nor false alarm exclusionary criteria were preregistered but applied for uniformity across experiments. Like Experiment 1, participants were kicked from the experiment for failing an attention check, so this is not listed as a criterion.

Results

The data analytic approach was identical to Experiment 1. Late responses were infrequent (overall M = .01, SE = .001) and not counted as correct. As with Experiment 1, PM performance, false alarms, PM block ongoing task performance (accuracy and response times), and PM target recognition memory were separately submitted to a 2 (load: low vs. high) × 2 (reminder: reminder vs. no reminder) ANOVA. Descriptive statistics for these variables of interest can be found in Table 1. Exploratory analyses for perceived PM task importance (2 [Load] × 2 [Reminder] ANOVA) and reminder checking frequency (one-way ANOVA) were also conducted. Descriptive statistics for these variables can be found in Table 2. For interaction effects of reminder by load, we used Bonferroni-corrected p values set to .025 (not preregistered) for the two comparisons between the reminder condition and the no-reminder control under high load and low load. The only preregistered analyses for Experiment 2 were for PM performance and ongoing task performance. We had also originally preregistered an analysis for detection sensitivity (i.e., PM hits minus PM false alarms), but on further inspection realized this was not a valid dependent variable because the proportional scoring of PM hits (of eight) and false alarms to nontargets (of 196) are not comparable.

PM Performance

Proportion Correct. The analysis for PM performance (Figure 3; Table 1) revealed that PM was higher with reminders, *Reminder:* $F(1, 144) = 16.02$, $p < .001$, $\eta_p^2 = .100$, and with low cognitive load, *Load:* $F(1, 144) = 21.26$, $p < .001$, $\eta_p^2 = .129$. There was also an interaction between the two, *Reminder* × *Load:* $F(1, 144) = 6.83$, $p = .010$, $\eta_p^2 = .045$. This interaction reflects that reminders did not improve PM in the low load condition ($F < 1$), but they did in the high load condition, $F(1, 71) = 28.36$, $p < .001$, $\eta_p^2 = .285$.

False Alarms. False alarms (see Table 1) were not influenced by reminders, *Reminder:* $F(1, 144) = 3.35$, $p = .069$, $\eta_p^2 = .023$, or load, *Load:* $F(1, 144) = 1.21$, $p = .274$, $\eta_p^2 = .008$, and there was no interaction between the two (*Reminder* × *Load:* $F < 1$).

Ongoing Task Performance

Accuracy. Accuracy (see Table 1) was not influenced by reminders, *Reminder:* $F(1, 144) = 3.53$, $p = .062$, $\eta_p^2 = .024$, or load ($F < 1$), and there was no interaction between the two ($F < 1$).

Response Time. Response times (see Table 1) were not influenced by reminders, *Reminder:* $F(1, 144) = 3.51$, $p = .063$, $\eta_p^2 = .024$, but greater load slowed responses, *Load:* $F(1, 144) = 5.35$, $p = .022$, $\eta_p^2 = .036$. There was no interaction between the two, *Reminder* × *Load:* $F(1, 144) = 1.36$, $p = .246$, $\eta_p^2 = .009$.

Recognition Memory

Recognition memory (see Table 1) was not influenced by reminders (*Reminder:* $F < 1$) or load (*Load:* $F < 1$), and there was no interaction between the two (*Reminder* × *Load:* $F < 1$).

Exploratory Analyses

PM Importance. Perceived importance (see Table 2) was not influenced by reminders, *Reminder:* $F(1, 144) = 1.79$, $p = .183$, $\eta_p^2 = .012$, or load, *Load:* $F(1, 144) = 3.06$, $p = .082$, $\eta_p^2 = .021$, and there was no interaction between the two, *Reminder* × *Load:* $F(1, 144) = 3.59$, $p = .060$, $\eta_p^2 = .024$.

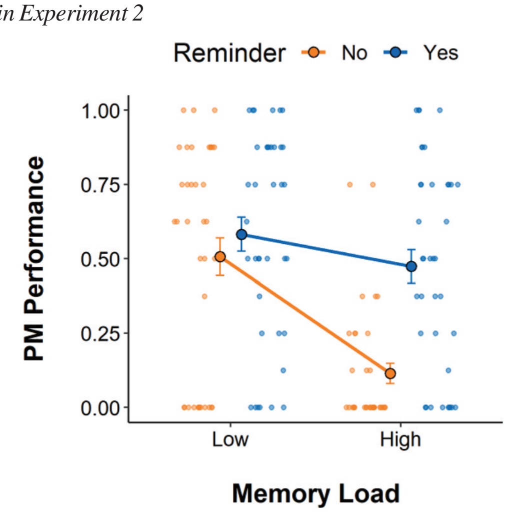
Reminder Checking. Reminder checking frequency (see Table 2) was greater under high load than low load, *Load:* $F(1, 74) = 29.34$, $p < .001$, $\eta_p^2 = .284$.

Discussion

Experiment 2 replicated the PM performance benefit of reminders under high cognitive load observed in Experiment 1, but with lower overall performance. The benefit to PM performance did not occur because of changes in recognition memory or increased perceived

Figure 3

Prospective Memory Performance as a Function of Reminder and Load in Experiment 2



Note. PM performance refers to the proportion of PM targets detected (out of eight). Large circles indicate mean performance, whereas the smaller circles reflect individual data points to illustrate the distribution of scores in each condition. Error bars reflect standard error. See the online article for the color version of this figure.

importance of the PM task but could be in part explained by the increased reminder checking in the high load reminder condition. There was also no overall effect of reminders on ongoing task performance, suggesting reminders did not increase monitoring for targets. However, Experiment 2 did find that load slowed ongoing task response times, which is consistent with the idea that nonspecific targets have a higher cognitive load and require more monitoring than specific targets (e.g., Smith, 2003). Because nonspecific reminders cannot enhance the discrepancy of targets and did not increase monitoring, this suggests that they may have improved PM performance by offloading the verification process. That is, participants could have used the nonspecific reminders to verify that a target belonged to a category relevant to the intention, which would be particularly helpful when having to remember multiple target categories in the high load condition.

Experiment 3

The first two experiments were simple intentions that required only a single action for each PM target. However, intentions in the real world are often more complex and require a unique action for each target (e.g., take one heart medication pill and two cholesterol pills). Complex intentions (i.e., unique actions) have greater retrieval demands than simple intentions (i.e., single action) and allow us to isolate and vary retrieval difficulty. Experiment 3 used complex intentions with a high target load (i.e., four targets), while varying the action load by manipulating the association between the target and the unique action paired with it. The association between these pairs influences the demand of retrieving the intended action.

Memory load in Experiment 3 was manipulated within-subjects by having participants learn two strongly associated pairs and two weakly associated pairs. The reminder condition had all four paired associates listed at the top of the screen (see Figure 1). Participants could use the target-action reminder to confirm the action rather than relying on a controlled search of memory, so we anticipated that reminders would produce a greater benefit to PM performance under high than low action retrieval load. Although specific reminders can facilitate recognition by increasing the discrepancy signal (as with the specific targets in Experiment 1) or increasing monitoring, this should occur for both strong associations and weak associations. Instead, Guynn et al. (1998) argued that reminders that include both target and action information strengthen the association between the between the two. In the current experiment, checking the target-action reminder could strengthen the association and enhance reflexive-retrieval of the intention. Thus, any benefits seen for PM under high memory load should occur because reminders reduce the need to engage in a difficult internal memory search and lead to reflexive-retrieval of the associated action. Experiment 3 preregistration information can be found at https://osf.io/fbt93 and data can be found at https://osf.io/z9sne (Peper et al., 2022b).[4]

Method

Participants

A power analysis for a 2 × 2 mixed-methods ANOVA based on the interaction effect size from Experiment 2 recommended 80 participants. A conservative sample of 100 participants was chosen to account for potential exclusionary data. The entire sample consisted of undergraduate students at the University of Texas at Arlington (age 18–28 years, M = 18.81, SD = 1.77) who received class credit for participation. Of the total 100 participants collected, 13 were excluded for meeting exclusionary criteria detailed below, for a final sample of 87. Participants were randomly assigned to either the reminder condition (N = 44) or the no reminder condition (N = 43).

Design

A 2 (load: low vs. high; *within*) × 2 (reminder: reminder vs. no reminder; *between*) mixed-method design was used. Load was manipulated within-subjects by varying the association between the target-action paired associates. All participants experienced both high (low-load) and low-association (high-load) target-action paired associates.

Materials

Ongoing task stimuli were the same as Experiment 1. For the PM target-action pairs, paired associates were chosen using the Florida Association Norms (Nelson et al., 2004). Strong association pairs (low load) had forward and backward associations both above .5 (e.g., *positive* – *negative* and *forward* – *backward*). Weak association (high load) pairs had no forward or backward association (e.g., *renounce* – *sculpture* and *conference* – *intimate*).

Procedure

The procedure was similar to those used in Experiment 1 and 2. It included an ongoing task practice with feedback (20 trials), ongoing task practice with feedback (40 trials), PM task instructions, PM task instructions quiz, PM task practice, PM target-action pair study, PM target-action pair quiz, distractor, PM block (200 trials), and post experimental questions.

There were a few changes to accommodate the difference in the PM response (i.e., typing the target word). The first change was to the ongoing task more generally. Rather than the ongoing task response resulting in the onset of the next trial, a "+" symbol appeared indicating that they were to press the spacebar to continue to the next trial (Cook et al., 2014). After practicing the ongoing task, participants received instructions for the PM task. Participants were instructed that they would learn four target-action paired associates and should type in the second word of the pair if they encountered the first word during the syllable judgment task. Specifically, participants were instructed to type in the second word during the "+" message that followed the ongoing task response. For example, when *positive* was presented, participants should have first pressed the "J" key to indicate it was a three-syllable word, then typed in the word "*negative*" when the "+" message appeared, and then pressed the spacebar to move onto the next trial. Participants were instructed to type "*idk*" if they noticed the PM target but forgot what the action word associate was.

Because this task is slightly more complicated than the previous procedures, a practice PM task was inserted after the PM task

[4] As described in footnote 2, the text of the preregistration document says "Experiment 4" although it corresponds to Experiment 3 in the current article.

instructions quiz (but prior to the PM target-action pair learning phase) to ensure the participants fully understood the instructions. Participants studied a practice target-action pair (*flower – sunlight*) and then completed a brief (5-trial) version of a PM block. Participants completed four syllable judgment trials before being presented with the PM cue (*flower*). Participants had to correctly type in the practice target word (*sunlight*) during the "+" message after making their ongoing task response to "*flower"* to continue to the next part of the study. If they did not do it correctly, they were brought back to the PM instructions screen and then redid the practice PM block. On successful completion of the practice, participants then studied the four target-action pairs for five seconds each. They were then quizzed on target-action pair using a two-alternative choice test as in Experiments 1 and 2. In the reminder condition, participants were instructed that the target-action pairs would be presented at the top of the screen (see Figure 1). Participants then completed the distractor task, PM block, and the same postexperimental questionnaires as Experiments 1 and 2. The recognition test was the same format as that used during the encoding phase, whereby one studied and one unstudied word pair were presented, and participants were instructed to select which pair they learned previously.

Analysis and Exclusionary Criteria

Participants were excluded for meeting the following preregistered criteria: (a) failing to have any PM hits *and* failing to recall the PM action (i.e., type the second word in the target-action pair) during the postexperimental questionnaire ($n = 4$); (b) reporting cheating (writing down PM target-action pairs; $n = 3$); (c) performing below 50% on the math distractor ($n = 2$); (d) achieving less than 75% ongoing task accuracy in the PM block ($n = 2$); (e) falling outside three standard deviations of the sample mean in ongoing task response times ($n = 2$); (f) false alarms ($n = 0$). This response time criterion was added to the preregistration to capture participants not putting in a good faith effort into the task. As in the previous experiments, participants were kicked out of the experiment and the data were not recorded if they failed an attention check.

Results

PM performance was submitted to a 2 (load: low vs. high; *within*) × 2 (reminder: reminder vs. no reminder; *between*) mixed method ANOVA. PM performance refers to targets (e.g., *positive*) correctly responded to with the action word (e.g., *negative*). PM block ongoing task performance (accuracy and response times) and PM target recognition memory were separately submitted and one-way ANOVAs. Descriptive statistics for these variables can be found in Table 1. For interaction effects of reminder by load, we used Bonferroni-corrected p values set to .025 for the two comparisons between the reminder condition and the no-reminder control under high load and low load. PM performance is a strict scoring method and refers to targets (e.g., *positive*) correctly responded to with the action word (e.g., *negative*). PM target noticing is a lenient scoring method that refers to participants responding to targets (e.g., *positive*) correctly *or* responding with "idk" or an incorrect word that indicated they noticed the target but failed to retrieve the correct action. False alarms were the number of action word, "idk," or incorrect word responses to nontarget trials. There were no late responses. An exploratory analysis for perceived PM task importance also applied a one-way ANOVA. Descriptive statistics for importance can be found in Table 2. The analyses for PM performance, target noticing, ongoing task performance, and false alarms were preregistered.

PM Performance

Proportion Correct. The analysis for PM performance (Figure 4; Table 1) revealed that PM was higher with reminders, *Reminder:* $F(1, 85) = 21.90$, $p < .001$, $\eta_p^2 = .205$, and with low load, *Load:* $F(1, 85) = 38.43$, $p < .001$, $\eta_p^2 = .311$. There was also an interaction between the two, *Reminder × Load:* $F(1, 85) = 9.16$, $p = .003$, $\eta_p^2 = .097$. This interaction reflects that although reminders improved PM under the low load, $F(1, 85) = 7.36$, $p = .008$, $\eta_p^2 = .080$, reminders improved PM to a greater extent under high load, $F(1, 85) = 32.56$, $p < .001$, $\eta_p^2 = .277$.

Target Noticing. Targets were noticed more often with reminders, *Reminder:* $F(1, 85) = 21.36$, $p < .001$, $\eta_p^2 = .201$, and with low load, *Load:* $F(1, 85) = 31.18$, $p < .001$, $\eta_p^2 = .268$. There was also an interaction between the two, *Reminder × Load:* $F(1, 85) = 5.73$, $p = .019$, $\eta_p^2 = .063$. This interaction reflects that although reminders improved target noticing under the low load, $F(1, 85) = 9.59$, $p = .003$, $\eta_p^2 = .101$, reminders improved target noticing to a greater extent under high load, $F(1, 85) = 27.57$, $p < .001$, $\eta_p^2 = .245$. As the pattern of target noticing mirrors that of PM performance, it will not be discussed further.

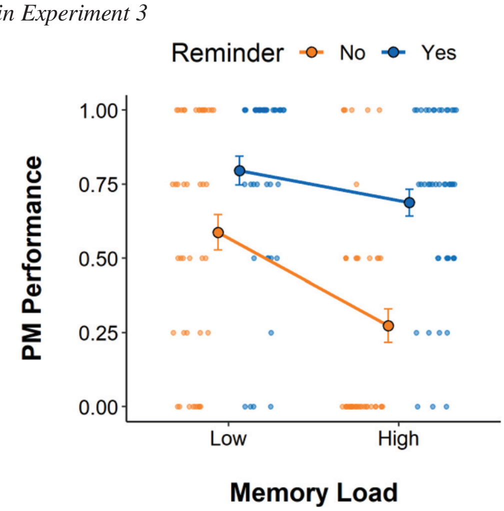
False Alarms. Reminders reduced the frequency of false alarms (see Table 1), Reminder: $F(1, 85) = 6.67$, $p = .012$, $\eta_p^2 = .073$.

Ongoing Task Performance

Accuracy. There was no effect of reminder (see Table 1), *Reminder:* $F(1, 85) = 1.32$, $p = .253$, $\eta_p^2 = .015$.

Figure 4

Prospective Memory Performance as a Function of Reminder and Load in Experiment 3



Note. Prospective memory (PM) performance refers to the proportion of PM targets responded to with the associated action (out of eight). Large circles indicate mean performance, whereas the smaller circles reflect individual data points to illustrate the distribution of scores in each condition. Error bars reflect standard error. See the online article for the color version of this figure.

Response Time. There was no effect of reminder (see Table 1), *Reminder:* $F(1, 85) = 2.81$, $p = .097$, $\eta_p^2 = .032$.

Recognition Memory

There was no effect of reminder (see Table 1; *Reminder:* $F < 1$).

Exploratory Analyses

For PM importance, there was no effect of reminder (see Table 2; *Reminder:* $F < 1$).

Discussion

Consistent with previous research, PM performance was significantly better under low than under high memory load (Cook et al., 2014; McDaniel et al., 2004). Critically, Experiment 3 replicated the results from Experiments 1 and 2 in that reminders provided a greater benefit to PM performance under high load. Additionally, the reminder benefit did not occur due to changes in ongoing task performance, recognition memory, or perceived PM task importance. These findings provide clear evidence that reminders are particularly beneficial when memory load is high.

Specific reminders for both the target and the action could have improved PM through two mechanisms. Target-action reminders may have facilitated target recognition (via enhancing target discrepancy or increased monitoring) or reduced demands on retrieval of the associated action following recognition. However, because memory load was manipulated within-subjects, it is not clear why target recognition should differ between the two. Additionally, reminders did not influence ongoing task performance, which rules out monitoring. The alternative view is that checking the reminders strengthened the association between the target and action (Guynn et al., 1998), which should primarily be beneficial for items with weak preexperimental associations. Thus, reminders presumably reduce demands on internal memory search processes by making retrieval more reflexive. To further isolate the two mechanisms (i.e., recognition and retrieval), Experiment 4 varied the type of reminder (target, action, or both).

Experiment 4

Although the previous experiments focused on memory load and reminders, they did not vary the type of reminder used. In the real world, people can set reminders in a variety of ways, but all reminders may not benefit PM equally. Experiment 4 manipulated the type of reminder (see Figure 1) using target-action word pairs with weak associations similar to Experiment 3 (e.g., *candle—elbow*) to assess the efficacy of reminders that promoted noticing the PM target (i.e., *candle;* target reminder), retrieving the PM action (i.e., *elbow*; action reminder), or both (i.e., *candle—elbow*; target-action reminder). Guynn et al. (1998) found that target-action and action-only reminders improved PM performance above that of a no-reminder control, but target-only reminders did not (however, see Einstein et al., 1998, for action-only reminders failing to improve time-based PM).

The purpose of the present experiment was to test the efficacy of different reminder types that work through separate mechanisms. Target reminders should primarily facilitate target recognition (i.e., increased fluency), action reminders should facilitate retrieval of the associated action, and target-action reminders should facilitate both. We predicted that all reminders would improve PM, but target-action reminders would benefit PM performance the most. We also predicted that target reminders would improve PM performance above that of action reminders, because noticing the target is a prerequisite for retrieving the associated action. Experiment 4 preregistration information can be found at https://osf.io/aqfg3 and data can be found at https://osf.io/t9bmg (Peper et al., 2022a).

Method

Participants

A power analysis for a one-way between-subjects ANOVA based on the effect of reminders under low load from the previous experiment ($\eta_p^2 = .080$) recommended 162 participants. A conservative sample of 200 participants was chosen to account for potential exclusionary data. Participants consisted of 202 undergraduate students at the University of Texas at Arlington (age 18–43, $M = 20.24$, $SD = 3.37$) who received class credit for participation. Thirty-one were excluded for meeting exclusionary criteria detailed below, for a final high load sample of 171. Participants were randomly assigned to the target reminder condition ($N = 41$), action reminder condition ($N = 43$), target-action reminder condition ($N = 45$), or the no reminder condition ($N = 42$).

Design, Materials, and Procedure

A four-level (reminder type: target-action reminder, target reminder, action reminder, or no reminder) between-subjects design was used. All participants experienced weak-association (i.e., high load; *candle—elbow*) target-action paired associates and one reminder type.

Materials

Ongoing task stimuli were the same as Experiment 1. For the PM target-action pairs, paired associates were chosen using the Florida Association Norms (Nelson et al., 2004). Weak association (high load) pairs had no forward or backward association (e.g., *renounce – sculpture* and *conference – intimate*).

Procedure

The procedure was similar to Experiment 3 except that all four target-action word pairs were weak-association, with each target word appearing twice in the PM task. After learning the target-action word pairs and being quizzed, participants in the target-action reminder condition were instructed that the target-action pairs would be presented at the top of the screen (see Figure 1). Participants in the target reminder condition were instructed that the target word would be listed at the top of the screen, and participants in the action reminder condition that the action word would be listed at the top of the screen.

Analysis and Exclusionary Criteria

Participants were excluded for meeting the following preregistered criteria: (a) failing to have any PM hits *and* failing to recall the PM action (i.e., type the second word in the target-action pair) during the postexperimental questionnaire ($n = 15$); (b) reporting cheating (writing down PM cues; $n = 4$); (c) performing below 50% on the math distractor ($n = 2$); (d) achieving less than 60% in ongoing task accuracy in the PM block ($n = 4$); (e) falling outside three standard deviations of the sample mean in ongoing task response times ($n = 3$); (f) duplicates ($n = 3$); (g) false alarms ($n = 2$). As in the previous experiments, participants were kicked out of the experiment and the data were not recorded if they failed an attention check.

Results

PM performance, PM target noticing, PM block ongoing task performance (accuracy and response times), PM target-action pair recognition memory, false alarms, and PM importance were submitted to a one-way (reminder type: target-action reminder, target reminder, action reminder, vs. no reminder) between-subjects ANOVA with four levels. Reminder checking frequency was analyzed with a one-way (reminder type: target-action reminder, target reminder, and action reminder) between-subjects ANOVA with three levels, because those in the no reminder conditions had no reminder to check and were therefore omitted from the analyses. Descriptive statistics for these variables are found in Table 1. For probing main effects of reminder type, we used Bonferroni-corrected p values set to .0167 for the three comparisons between each of the three reminder types with the no-reminder control. PM performance, target noticing, and false alarms were calculated in the same manner as Experiment 3. The analyses for perceived PM task importance and reminder checking frequency were exploratory. Descriptive statistics for these variables are found in Table 2. Late responses were infrequent (overall $M = .001$, $SE < .001$) and not counted as correct. Descriptive statistics for the primary variables can be found in Table 1, and exploratory variables can be found in Table 2. The analyses for PM performance, target noticing, ongoing task performance, and recognition memory were preregistered.

PM Performance

Proportion Correct. The analysis for PM performance (Figure 5; Table 1) revealed a main effect of reminders, *Reminder:* $F(3, 165) = 17.21$, $p < .001$, $\eta_p^2 = .238$. Compared with the no-reminder control, target-action reminders, $F(1, 83) = 34.31$, $p < .001$, $\eta_p^2 = .292$, improved performance, while target reminders ($F < 1$) and action reminders, $F(1, 82) = 5.83$, $p = .024$, $\eta_p^2 = .060$, did not after Bonferroni-corrections.

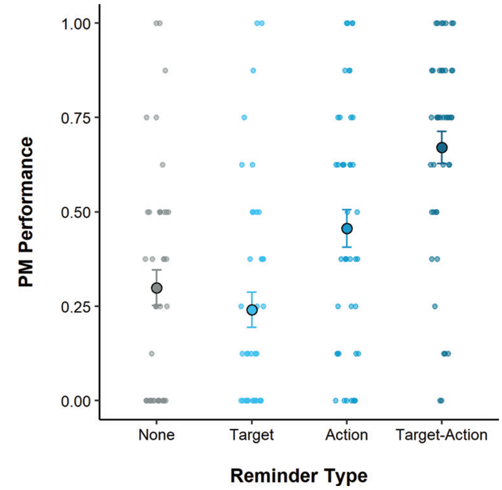
Target Noticing. The analysis for target noticing revealed a main effect of reminders, *Reminder:* $F(3, 165) = 15.78$, $p < .001$, $\eta_p^2 = .223$. Compared with the no-reminder control, target-action reminders improved performance, $F(1, 83) = 33.02$, $p < .001$, $\eta_p^2 = .285$, whereas target reminders ($F < 1$) and action reminders, $F(1, 82) = 4.38$, $p = .040$, $\eta_p^2 = .051$, did not after Bonferroni corrections.

False Alarms. There was no effect of reminder type (see Table 1), *Reminder:* $F(3, 165) = 2.36$, $p = .074$, $\eta_p^2 = .041$.

Ongoing Task Performance

Accuracy and Response Times. Reminder type had no effect on accuracy (see Table 1; *Reminder:* $F < 1$) or response times (*Reminder:* $F < 1$).

Figure 5
Prospective Memory Performance as a Function of Reminder Type in Experiment 4



Note. Prospective memory (PM) performance refers to the proportion of PM targets responded to with the associated action (out of eight). Large circles indicate mean performance, whereas the smaller circles reflect individual data points to illustrate the distribution of scores in each condition. Error bars reflect standard error. See the online article for the color version of this figure.

Recognition Memory. There was no effect of reminder type (see Table 1), *Reminder:* $F(3, 165) = 2.19$, $p = .092$, $\eta_p^2 = .038$.

Exploratory Analyses

PM Importance. There was no effect of reminder type (see Table 2; *Reminder:* $F < 1$).

Reminder Checking. For the three reminder conditions, there was no effect of reminder type on checking frequency (see Table 2), *Reminder:* $F(3, 165) = 1.01$, $p = .367$, $\eta_p^2 = .016$.

Discussion

Consistent with Experiment 3, target-action reminders improved PM performance under high memory load (i.e., weak target-action associations). Contrary to our predictions, neither target-only nor action-only reminders improved PM performance. These findings are consistent with the idea that target-action reminders strengthen the association between the weakly associated PM targets and action, making intention retrieval more reflexive (Guynn et al., 1998). However, it appears that when using complex intentions, increasing the discrepancy of the targets (target-only reminder) or reducing action retrieval demands (action-only reminder) alone are insufficient to improve PM performance. According to the associative view (Einstein & McDaniel, 1996; McDaniel, 1995), PM intentions reside at a high level of activation when the link between the target and action is sufficiently strong. Without a strong link between target and action, participants may not recognize the significance of a PM target (i.e., as a cue for the intention), let alone retrieve the associated action. Participants reported checking the different reminder types equally across all conditions, suggesting people are not aware of differences in reminder effectiveness (or lack thereof). Additionally, the benefit of reminders again was not due to differences in recognition memory or perceived importance of the PM task. Finally, there was no influence of reminders on ongoing task performance. These findings suggest that not all reminders are created equal, and that the most effective reminders for complex intentions include both the PM target and the associated action.

General Discussion

The present study aimed to understand the processes underlying how reminders influence PM performance under different memory loads. Target reminders improved PM for simple intentions (Experiments 1 and 2). For complex intentions, target-action reminders improved PM (Experiments 3 and 4), whereas target or action reminders alone did not (Experiment 4). Importantly, the improvements in PM from reminders were greatest when there was high memory load and occurred without any cost to ongoing task performance or reduction of retrospective memory for the targets at the end of the experiment. However, not all reminder types may be equally effective. Below we discuss the theoretical and applied ramifications of such findings.

Previous research has shown that increased cognitive load—whether that be manipulated through target specificity, the number of targets, or retrieval demands—can negatively impact PM performance (Cook et al., 2014; Einstein et al., 2005; Marsh et al., 2003; McDaniel et al., 2004; Wesslein et al., 2014). This pattern was replicated across the first three experiments of the current study. That is, without the use of reminders, PM performance was better for single target compared with multitarget intentions in Experiments 1 and 2 and for strongly associated relative to weakly associated pairs in Experiment 3. These results are consistent with predictions from the Multiprocess Framework which posits that task parameters determine the extent to which effortful control processes are used to retrieve PM intentions (Einstein et al., 2005). Importantly, we also replicate previous studies showing that compensatory strategies such as offloading can be used to improve PM, particularly under high demand (Chen et al., 2017; Cherkaoui & Gilbert, 2017; Gilbert, 2015a, 2015b; Gilbert et al., 2020; Guynn et al., 1998; Henry et al., 2012; Loft et al., 2011; Vortac et al., 1995).

Microstructure of Prospective Memory

Based on prior theorizing about the microprocesses associated with prospective remembering (e.g., Marsh et al., 2003), we reasoned that reminders may change demands associated with recognition, verification, and/or retrieval of the intention. Using specific target reminders (Experiments 1, 3, and 4) presumably increases the fluency of PM targets, such that each time the reminder is checked, the target discrepancy signal is strengthened. It follows that reminders were more useful with more targets (i.e., high load), as a single target may elicit a discrepancy signal even without reminders. Similarly, Guynn and McDaniel (2007) exposed participants to PM targets prior to encoding. Preexposure to targets improved PM performance due to changing the discrepancy of targets by increasing target familiarity. Notably, specific reminders produced a small PM improvement under low load in Experiments 1 and 3, further highlighting their utility, and suggesting that reminders may still increase the discrepancy signal even for a single target (Morita, 2006).

In contrast, when nonspecific targets were used in Experiment 2, there was no reminder effect under low load, consistent with the idea that nonspecific reminders cannot influence increase the discrepancy signal for specific category exemplar PM targets. However, in this case, participants must verify whether the current stimulus (e.g., *parakeet*) matches one of the nonspecific categories maintained in working memory (e.g., *bird*), which is more demanding with multiple categories. Thus, reminders can be used to offload the maintenance of the category representations, thereby facilitating the verification (matching) process of an exemplar with the target category.

Finally, reminders may facilitate retrieval processes, at least when different actions are required for different targets (Experiments 3 and 4). Each time a reminder is checked, the association between the target and action may be strengthened (Guynn et al., 1998), increasing the likelihood of reflexive retrieval of the intended action. As shown in the current study, this should be particularly beneficial for pairs with weak preexperimental associations. Target-only and action-only reminders were not sufficient for improving intention retrieval (Experiment 4), suggesting that reminders that improve both recognition and retrieval processes are important for remembering. Importantly, the fact that reminders did not significantly influence ongoing task performance in any experiment rules out the possibility reminders resulted in greater monitoring. Together these findings highlight that different reminders may influence different processes associated with PM and that the efficacy of these reminders depends on the nature of the reminder and PM task.

Although the primary focus of the current study was on successful PM retrieval, we also examined whether reminders influenced false alarms to nontarget trials. Loft et al. (2011) found that reminders presented briefly before a target appeared reduced false alarms. However, there was no effect when the reminders were presented throughout the entire task. Although it is not entirely clear why there were differences in false alarm rates between periodic and always-available reminders, it should be noted that always-available reminders also did not aid PM performance—unlike the current study and previous research (Chen et al., 2017; Einstein & McDaniel, 1990; Henry et al., 2012; Vortac et al., 1995). In the current study, two out of the three experiments (Experiments 1 and 3) using specific targets reminders resulted in reduced false alarms to nontarget stimuli. Participants could have used the specific reminders to verify the stimulus as a nontarget when they were unsure. This implies that specific reminders could facilitate memory via target recognition and action retrieval, while also reducing false remembering by facilitating verification. These findings highlight another utility of reminders by showing, at least in some instances, they can reduce errors. Examining the role of reminders in other paradigms (e.g., output monitoring, commission error) may provide valuable information about how to reduce errors that can have important real-world consequences (e.g., over- or undermedication).

One final point of interest is that the greater reminding checking under high load did not influence ongoing task performance. One possibility is that reminders reduced monitoring demands, and the additional time taken to check the reminder washed out any response time differences. However, this does not explain differences (or lack thereof) between load conditions. Another possibility is that participants checked reminders briefly during the intertrial interval and/or did not check frequently enough to significantly influence response times across the entire PM task. In the latter case, it is possible that even minimal reminder checking could enhance target memory enough to improve PM performance.

When Reminders Are Helpful

Previous research has found that sometimes target reminders are sufficient for improving PM (e.g., Vortac et al., 1995), whereas other times both target and actions are required (Guynn et al., 1998). The current study showed that target reminders were sufficient when participants learned a single action (i.e., press the "7" key), but that both target and action reminders were necessary to improve PM when participants learned unique actions associated with each target. It follows then that target reminders may be sufficient when the action retrieval demands are low, but target-action reminders are needed when action retrieval demands are high (McDaniel et al., 2004). This would suggest that one potential reason why prior research using the intention offloading task has found reminders to be so successful for improving PM is because they specify the target (e.g., the "G" circle) *and* the action (e.g., drag to upper red location). These reminders may be considerably less effective if only the target information was specified.

Another important thing to consider are the metacognitive processes involved in PM reminders. Previous research using the intention offloading task has allowed participants to choose whether they want to set reminders (e.g., Gilbert, 2015a; Gilbert et al., 2020). These studies have shown that metacognitive biases can limit the utility of reminders, as being overconfident in one's internal memory ability results in lower reminder usage. In the current study, presenting reminders on the screen for participants limits these potential biases. Indeed, many studies using more traditional PM paradigms that show benefits from reminders present reminders for participants (i.e., are experimenter-generated) and have the reminders available throughout the entire task. Critically, however, although experimenter-generated reminders remove the metacognitive control required to strategically *set* reminders, metacognitive control is still involved in the decision to *check* the reminder. In Experiments 1 and 2, participants reported checking the reminder more frequently under high memory load, suggesting that participants were well-calibrated with task demands. However, participants may be less calibrated with differences in the efficacy of various reminder types. Participants in Experiment 4 reported checking reminders with equal frequency across the three reminder conditions, despite only target-action reminders producing benefits. In this case, expecting to have reminders, but not realizing that target-only or action-only reminders are ineffective, may result in overreliance on reminders that do not help PM performance. Thus, metacognitive processes must still be taken into consideration when understanding when and how reminders will be most efficacious.

The current also study explored whether having reminders influenced later memory of the targets during a postexperimental recognition test. Prior research in the retrospective memory domain has examined the influence of expected reminders on memory recall. For example, Kelly and Risko (2019) had participants engage in multiple study-test opportunities with reminders. On the final test, the reminder was unexpectedly taken away. Those who expected a reminder had worse memory than those in a condition who were aware they would not have a reminder. In the present study, participants had a reminder for the PM task, but did not have a reminder when their recognition memory was tested at the end of the procedure. However, this did not appear to negatively impact memory, because recognition was quite high in both conditions across all experiments. It is possible that the difference in findings compared with previous research is attributable to the greater exposure to the PM targets (and their associated actions) that could have provided further opportunities for rehearsing and encoding the targets. Alternatively, these findings may simply reflect that performance was near ceiling given so few targets (and actions) were learned in all experiments. This lack of variability in recognition memory means that the absence of differences across conditions should interpreted with caution (as similarly noted for the high PM performance under low load seen in Experiment 1). However, our primary takeaway from the fact that recognition memory was high in all conditions means that PM target information (along with action information in Experiments 3 and 4) was indeed available in memory when participants were explicitly put in a retrospective retrieval mode (i.e., cued retrieval). However, when participants were in a *prospective* retrieval mode that requires self-initiated retrieval, participants clearly had difficulty *accessing* this information while also performing the demanding ongoing task. These findings suggest that reminders may increase accessibility of PM-relevant information when self-initiated retrieval is required, even when the contents of the intention are available in memory.

Alternative Account

Although we have interpreted these findings based on the Multiprocess Framework and the microstructure accounts of PM, these results may also be explained by the Preparatory Attentional and Memory processes theory (Smith, 2003). This theory states attentionally demanding preparatory processes are needed to actively search the environment targets and stimulates a recognition check to determine whether the current stimulus is appropriate for making a PM response. This theory has been formalized using multinomial modeling that when fit to PM data produces two parameter values: an attention parameter (i.e., remembering *that* something is to be done) and a memory parameter (i.e., remembering *what* is to be done; Smith & Bayen, 2004, 2005). Previous research has found that reducing target specificity produces changes in the attention, whereas increasing the number of targets changes both the attention and memory parameters (Wesslein et al., 2014). In the current study, the finding that reminders (independent of load) benefited both specific (Experiments 1 and 3) and nonspecific (Experiment 2) targets in the current study suggests that reminders may reduce demands on attention. This is further supported by the small benefit of reminders for specific targets even under low load when memory demands should be negligible. Reminders likely also lessened demands on the memory component in Experiments 1 and 2, as the benefit of reminders became much stronger as the number of targets increased. Furthermore, the direct manipulation of memory demands in Experiment 3 adds clearer support for the idea that reminders improved the memory component of PM. Future research could use multinomial modeling and specifically test how reminders influence the memory and attention components of PM.

Conclusions

Four experiments demonstrated PM is improved by reminders of various types under different circumstances, especially when the PM load is high. However, not all reminder types are beneficial. There are numerous ways to vary reminders (and cognitive load), but we believe this paradigm will be fruitful for exploring many questions regarding the influence of reminders on laboratory PM. In everyday life, effective PM ability is essential for navigating a complex world and maintaining independence with increased age (Woods et al., 2012). Remembering and executing the multitudes of daily intentions (e.g., appointments, medications, grocery lists, etc.) and various spontaneous tasks that arise from moment to moment (e.g., replying to emails, getting gas, filling a pet's water bowl, etc.) can be difficult. Given the current findings, reminders provide a promising tool for supporting populations with low PM abilities (e.g., schizophrenia) or general cognitive decline (e.g., older adults). The findings that reminders added no cost to ongoing task performance further highlight the utility of externally supporting PM.

References

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

Ball, B. H., & Bugg, J. M. (2018). Context cue focality influences strategic prospective memory monitoring. *Psychonomic Bulletin & Review*, *25*(4), 1405–1415. https://doi.org/10.3758/s13423-018-1442-9

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

Breneiser, J. E., & McDaniel, M. A. (2006). Discrepancy processes in prospective memory retrieval. *Psychonomic Bulletin & Review*, *13*(5), 837–841. https://doi.org/10.3758/BF03194006

Brewer, G. A., Knight, J. B., Marsh, R. L., & Unsworth, N. (2010). Individual differences in event-based prospective memory: Evidence for multiple processes supporting cue detection. *Memory & Cognition*, *38*(3), 304–311. https://doi.org/10.3758/MC.38.3.304

Bulley, A., McCarthy, T., Gilbert, S. J., Suddendorf, T., & Redshaw, J. (2020). Children devise and selectively use tools to offload cognition. *Current Biology*, *30*(17), 3457–3464.e3.

Chen, Y., Lian, R., Yang, L., Liu, J., & Meng, Y. (2017). Working memory load and reminder effect on event-based prospective memory of high-and low-achieving students in math. *Journal of Learning Disabilities*, *50*(5), 602–608. https://doi.org/10.1177/0022219416668322

Cherkaoui, M., & Gilbert, S. J. (2017). Strategic use of reminders in an 'intention offloading' task: Do individuals with autism spectrum conditions compensate for memory difficulties? *Neuropsychologia*, *97*, 140–151. https://doi.org/10.1016/j.neuropsychologia.2017.02.008

Cohen, A. L., Jaudas, A., & Gollwitzer, P. M. (2008). Number of cues influences the cost of remembering to remember. *Memory & Cognition*, *36*(1), 149–156. https://doi.org/10.3758/MC.36.1.149

Cook, C., Ball, B. H., & Brewer, G. A. (2014). No effects of executive control depletion on prospective memory retrieval processes. *Consciousness and Cognition*, *27*, 121–128. https://doi.org/10.1016/j.concog.2014.04.006

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

Einstein, G. O., & McDaniel, M. A. (1996). Retrieval processes in prospective memory: Theoretical approaches and some new empirical findings. In M. Brandimonte, G. O. Einstein & M. A. McDaniel (Eds.), *Prospective Memory: Theory and Applications* (pp. 115–141). Psychology Press.

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science*, *14*(6), 286–290. https://doi.org/10.1111/j.0963-7214.2005.00382.x

Einstein, G. O., & McDaniel, M. A. (2010). Prospective memory and what costs do not reveal about retrieval processes: A commentary on Smith, Hunt, McVay, and McConnell (2007). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(4), 1082–1088. https://doi.org/10.1037/a0019184

Einstein, G. O., McDaniel, M. A., Smith, R. E., & Shaw, P. (1998). Habitual prospective memory and aging: Remembering intentions and forgetting actions. *Psychological Science*, *9*(4), 284–288. https://doi.org/10.1111/1467-9280.00056

Einstein, G. O., McDaniel, M. A., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., & Breneiser, J. (2005). Multiple processes in prospective memory retrieval: Factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology: General*, *134*(3), 327–342. https://doi.org/10.1037/0096-3445.134.3.327

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

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

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

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

Guynn, M. J., & McDaniel, M. A. (2007). Target preexposure eliminates the effect of distraction on event-based prospective memory. *Psychonomic Bulletin & Review*, *14*(3), 484–488. https://doi.org/10.3758/BF03194094

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

Heathcote, A., Loft, S., & Remington, R. W. (2015). Slow down and remember to remember! A delay theory of prospective memory costs. *Psychological Review*, *122*(2), 376–410. https://doi.org/10.1037/a0038952

Henry, J. D., Rendell, P. G., Phillips, L. H., Dunlop, L., & Kliegel, M. (2012). Prospective memory reminders: A laboratory investigation of initiation source and age effects. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *65*(7), 1274–1287. https://doi.org/10.1080/17470218.2011.651091

Hicks, J. L., Franks, B. A., & Spitler, S. N. (2017). Prior task experience and comparable stimulus exposure nullify focal and nonfocal prospective memory retrieval differences. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *70*(10), 1997–2006. https://doi.org/10.1080/17470218.2016.1217891

Humphreys, M. S., Li, Y. R., Burt, J. S., & Loft, S. (2020). How semantic processing affects recognition memory. *Journal of Memory and Language*, *113*, 104109.

Ihle, A., Schnitzspahn, K., Rendell, P. G., Luong, C., & Kliegel, M. (2012). Age benefits in everyday prospective memory: The influence of personal task importance, use of reminders and everyday stress. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *19*(1–2), 84–101. https://doi.org/10.1080/13825585.2011.629288

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

Loft, S., Smith, R. E., & Bhaskara, A. (2011). Prospective memory in an air traffic control simulation: External aids that signal when to act. *Journal of Experimental Psychology: Applied*, *17*(1), 60–70. https://doi.org/10.1037/a0022845

Marsh, R. L., Hicks, J. L., & Watson, V. (2002). The dynamics of intention retrieval and coordination of action in event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(4), 652–659. https://doi.org/10.1037/0278-7393.28.4.652

Marsh, R. L., Hicks, J. L., Cook, G. I., Hansen, J. S., & Pallos, A. L. (2003). Interference to ongoing activities covaries with the characteristics of an event-based intention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 861–870. https://doi.org/10.1037/0278-7393.29.5.861

McDaniel, M. A. (1995). Prospective memory: Progress and processes. *Psychology of Learning and Motivation*, *33*, 191–222. https://doi.org/10.1016/S0079-7421(08)60375-8

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field*. Sage. https://doi.org/10.4135/9781452225913

McDaniel, M. A., Guynn, M. J., Einstein, G. O., & Breneiser, J. (2004). Cue-focused and reflexive-associative processes in prospective memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(3), 605–614. https://doi.org/10.1037/0278-7393.30.3.605

Meier, B., & Rey-Mermet, A. (2012). Beyond monitoring: After-effects of responding to prospective memory targets. *Consciousness and Cognition*, *21*(4), 1644–1653. https://doi.org/10.1016/j.concog.2012.09.003

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

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

Peper, P., Alakbarova, D., & Ball, B. H. (2022a). Experiment 4 [Data set]. https://osf.io/ruwtv/

Peper, P., Alakbarova, D., & Ball, B. H. (2022b). Experiments 1, 2, and 3 [Data set]. https://osf.io/z59ru/

Redshaw, J., Vandersee, J., Bulley, A., & Gilbert, S. J. (2018). Development of children's use of external reminders for hard-to-remember intentions. *Child Development*, *89*(6), 2099–2108. https://doi.org/10.1111/cdev.13040

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

Schnitzspahn, K. M., Kvavilashvili, L., & Altgassen, M. (2020). Redefining the pattern of age-prospective memory-paradox: New insights on age effects in lab-based, naturalistic, and self-assigned tasks. *Psychological Research*, *84*(5), 1370–1386. https://doi.org/10.1007/s00426-018-1140-2

Scullin, M. K., McDaniel, M. A., Shelton, J. T., & Lee, J. H. (2010). Focal/nonfocal cue effects in prospective memory: Monitoring difficulty or different retrieval processes? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 736–749. https://doi.org/10.1037/a0018971

Shelton, J. T., Scullin, M. K., & Hacker, J. Y. (2019). The multiprocess framework: Historical context and the "dynamic" extension. In *Prospective memory* (pp. 10–26). Routledge. https://doi.org/10.4324/9781351000154-2

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

Smith, R. E., & Bayen, U. J. (2004). A multinomial model of event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(4), 756–777. https://doi.org/10.1037/0278-7393.30.4.756

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

Strickland, L., Heathcote, A., Humphreys, M. S., & Loft, S. (2022). Target learning in event-based prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(8), 1110–1126. https://doi.org/10.1037/xlm0000900

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the norms. *Journal of Memory and Language*, *50*(3), 289–335. https://doi.org/10.1016/j.jml.2003.10.003

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

Wesslein, A.-K., Rummel, J., & Boywitt, C. D. (2014). Differential effects of cue specificity and list length on the prospective and retrospective prospective-memory components. *Journal of Cognitive Psychology*, *26*(2), 135–146. https://doi.org/10.1080/20445911.2013.865628

West, R., & Craik, F. I. (1999). Age-related decline in prospective memory: The roles of cue accessibility and cue sensitivity. *Psychology and Aging*, *14*(2), 264–272.

Woods, S. P., Weinborn, M., Velnoweth, A., Rooney, A., & Bucks, R. S. (2012). Memory for intentions is uniquely associated with instrumental activities of daily living in healthy older adults. *Journal of the International Neuropsychological Society*, *18*(1), 134–138. https://doi.org/10.1017/S1355617711001263

Received March 4, 2021
Revision received August 24, 2022
Accepted August 28, 2022 ■