American Psychological Association

Journal of Experimental Psychology: Learning, Memory, and Cognition

© 2020 American Psychological Association
ISSN: 0278-7393

http://dx.doi.org/10.1037/xlm0000976

Wait a Second . . . Boundary Conditions on Delayed Responding Theories of Prospective Memory

B. Hunter Ball
University of Texas at Arlington

Anne Vogel
University of Mississippi

Derek M. Ellis and Gene A. Brewer
Arizona State University

Research suggests that forcing participants to withhold responding for as brief as 600 ms eliminates one of the most reliable findings in prospective memory (PM): the cue focality effect. This result undermines the conventional view that controlled attentional monitoring processes support PM, and instead suggests that cue detection results from increased response thresholds that allow more time for PM information to accumulate. Given the significance of such findings, it is critical to examine the generalizability of the delay mechanism. Experiments 1–4 examined boundary conditions of the delay theory of PM, whereas Experiment 5 more directly tested contrasting theoretical predictions from monitoring theory (e.g., multiprocess framework) and delay theory. Using the same (Experiment 1) or conceptually similar (Experiment 2) delay procedure and identical cues (nonfocal "tor" intention) from the original study failed to show any influence of delay on performance. Using a different nonfocal intention (first letter "S") similarly did not influence performance (Experiment 3), and the difference between focal and nonfocal cue detection was never completely eliminated even with delays as long as 2,500 ms (Experiment 4). Experiment 5 did find the anticipated reduction in the focality effect with increased delays with a larger sample ($n = 249$). However, the focality effect was not moderated by attention control ability despite the fact that participants with impoverished attention control should benefit most from the delay procedure. These results suggest that any theory of PM that considers only a delay mechanism may not fully capture the dynamic attention processes that support cue detection.

Keywords: prospective memory, attention, cue focality, individual differences

Supplemental materials: http://dx.doi.org/10.1037/xlm0000976.supp

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Our lives are filled with plans that often cannot be performed at the current moment. Remembering these plans at the appropriate moment in the future is referred to as *prospective memory* (PM). *Event-based PM* specifically refers to using environmental cues to trigger retrieval of delayed intentions. For example, encountering a medicine bottle on the bathroom counter may serve as a cue to take one's medication for the day. PM is unique from retrospective memory in that there is no explicit query of memory at the appropriate moment. Rather, successful noticing of PM cues and retrieval of intended actions must be self-initiated (Craik, 1986). Unfortunately, ongoing activities (e.g., household chores) often interfere with these processes, which can result in high rates of forgetting. Given the possible ramifications of PM failures (e.g., health issues due to undermedication), considerable research has been aimed at characterizing the theoretical mechanisms underlying PM to better understand how these failures can be reduced. The current study revisits these mechanisms in light of more recent theoretical developments that challenge traditional views of PM processing.

In a typical PM paradigm, participants are given an intention to make a special response to a particular set of cues while busily engaged in an ostensibly unrelated ongoing task. For example, participants may be instructed to press the "7" key (i.e., the PM action) any time the word *sister* (i.e., the PM cue) is encountered while deciding whether a string of letters form a word or a nonword (i.e., the ongoing task). One of the most reliable findings using this procedure is that PM performance (pressing the "7" key)

B. Hunter Ball, Department of Psychology, University of Texas at Arlington; Anne Vogel, Department of Psychology, University of Mississippi; Derek M. Ellis and Gene A. Brewer, Department of Psychology, Arizona State University.

This work was supported by National Science Foundation Grant 1632291 awarded to Gene A. Brewer. B. Hunter Ball was supported by an NIA Training Grant (T32AG000030-40) at Washington University during data collection for some portions of the manuscript. All data is available upon request.

We thank Daisy Estrada, Margaret Hargrave, Kierra Harris, Marissa Kepple, Christina Rao, Monica Rentowski, Ashree Subedi, and Erica Williams for their assistance in data collection. Portions of the data were presented the ARMADILLO Conference is Houston, Texas, and the Annual Meeting of the Psychonomics Society in New Orleans, Louisiana.

Correspondence concerning this article should be addressed to B. Hunter Ball, Department of Psychology, University of Texas at Arlington, 501 Nedderman Drive, Arlington, TX 76019. E-mail: Hunter.Ball@uta.edu

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

and ongoing task performance (speed and/or accuracy of lexical decisions) differs considerably depending on the nature of the PM cues embedded within the task. A PM cue is considered "focal" or "nonfocal," respectively, when ongoing task processing does or does not automatically orient attention to the relevant features of the PM cue (Einstein & McDaniel, 2005; Einstein et al., 2005). For example, during a lexical-decision task the specific word *sister* would be considered a focal cue, whereas any word starting with the letter "S" would be considered a nonfocal cue. This distinction occurs because accessing word information from memory for each item during a lexical-decision task orients attention to the specific word *sister* but does not orient attention to the first letter of each item. The focality effect refers to the finding that, almost invariably, focal cues are detected at higher rates than nonfocal cues. Relatedly, ongoing task performance is typically faster and/or more accurate in focal conditions than nonfocal conditions (Anderson, Strube, & McDaniel, 2019; Einstein et al., 2005). When comparing ongoing task performance in a focal condition to a control condition with no PM intention, cost to ongoing task performance due to possessing a focal intention is typically negligible. That is, the high rates of focal cue detection typically occur with little to no sacrifices in terms of speed or accuracy to ongoing task performance. In contrast, the considerably lower rate of nonfocal cue detection is accompanied by cost to ongoing task performance.

The most prominent account of these findings is the multiprocess framework (McDaniels & Einstein, 2000; see also Scullin, McDaniel, & Shelton, 2013, and Shelton & Scullin, 2017, for updates to this theory). This framework suggests that PM retrieval can occur via two processes: spontaneous retrieval and preparatory monitoring. Focal cue detection is thought to occur via spontaneous retrieval, which is a hippocampally mediated process that results in relatively automatic retrieval of the intention following associative cueing or discrepancy processing (Cona, Bisiacchi, Sartori, & Scarpazza, 2016; Gordon, Shelton, Bugg, McDaniel, & Head, 2011; McDaniel, Umanath, Einstein, & Waldum, 2015). In contrast, nonfocal cue detection is thought to occur via preparatory monitoring, which is a frontally mediated process that results in cue detection following engagement of attentionally demanding maintenance and active search of PM cues (Ball & Brewer, 2018; Brewer, Knight, Marsh, & Unsworth, 2010; Burgess, Quayle, & Frith, 2001; McDaniel, LaMontagne, Beck, Scullin, & Braver, 2013). Because monitoring for nonfocal cues requires capacity-consuming attention processes, this reduces processing resources available for ongoing task processing. With fewer resources available for ongoing task processing, performance suffers as evidenced by slower or less accurate ongoing task responding, referred to as *cost* (Smith, 2003). In contrast, spontaneous retrieval is relatively resource free, thereby producing negligible cost to ongoing task performance. Of course, the focality effect has been the source of contentious theoretical debate (see Einstein & McDaniel, 2010; Smith, 2010; Smith, Hunt, McVay, & McConnell, 2007). For example, the preparatory attention and memory (PAM) processes theory suggests that preparatory monitoring is always needed for successful cue detection, regardless of cue type (Smith, 2003). This contention mainly stems from disagreements of whether cost to ongoing task performance due to possessing an intention is ever completely eliminated in focal cue conditions. This debate aside, both theories agree that capacity-consuming attentional monitoring processes are needed to detect nonfocal cues.

More recent work, however, has challenged these prevailing views of PM processing. Using an experimental procedure similar to that described earlier, Loft and Remington (2013) cleverly instructed participants to withhold ongoing task responding until after a tone played. Tone onset ranged anywhere from 0 to 1,600 ms poststimulus onset. That is, either a tone was played immediately (0 ms) so participants could respond as soon as the lexical stimulus was presented (similar to standard procedures without a tone), or the participants had to withhold their ongoing task response until after a tone that was presented after various delays (e.g., 600 ms, 1,000, or 1,600 ms) following the onset of the lexical stimulus. Performance on the delay trials were compared to immediate trials (0 ms), the latter of which is how PM is traditionally assessed (i.e., with immediate responding). In one experiment using a specific word (e.g., offer) as the focal cue and a categorical intention (e.g., animals) as the nonfocal cue, it was found that the focality effect, as measured by the difference in focal and nonfocal cue detection, was completely eliminated with delays as short as 600 ms. In another experiment using a nonfocal syllable cue (e.g., tor), it was found that the focality effect was eliminated by having participants delay their response by 1,600 ms. The difference in duration to eliminate the focality effect across experiments likely reflects that syllable information is more difficult to notice than categorical information (Anderson, McDaniel, & Einstein, 2017; Anderson & McDaniel, 2019). In any manner, complete elimination of the focality effect is particularly astounding given that the difference between focal and nonfocal cue detection is consistently found in the literature. Based on these findings, it was suggested that it may take more time for nonfocal cue information to accumulate and compete for response selection than focal cue information. Thus, by forcing participants to withhold their responses for brief delays, this decreased the likelihood that the more routine ongoing task response would preempt PM retrieval on the relatively infrequent cue trials. This account suggests that if enough time is given for PM information to accumulate (e.g., 2,500 ms), then presumably nonfocal cue detection should always approximate focal cue detection and both should reach ceiling levels of performance in the limit. Extrapolating from these ideas, it was further suggested that the typical slowing in standard nonfocal tasks (without tones) relative to focal tasks might reflect that participants are endogenously implementing a similar delay process to support PM retrieval.

The ideas put forth by Loft and Remington (2013) have since been tested more formally by fitting evidence accumulation models to PM ongoing task data. Evidence accumulation models (e.g., drift diffusion, linear ballistic accumulator) simultaneously account for speed and accuracy of ongoing task decisions and allow for the dissociation of specific cognitive processes thought to contribute to the decision process (Brown & Heathcote, 2008; Ratcliff, 1978). The core of these processes includes the rate of information accumulation (drift rate), the amount of evidence required to make a decision (boundary separation), and peripheral processes occurring either prior to or after the actual decision (nondecision time). Heathcote, Loft, and Remington (2015) argued that if cue detection results from drawing attentional resources away from the ongoing task, as may be inferred from monitoring theories of PM, this should decrease rates of information accumulation (i.e., drift rates). That is, as more resources are devoted to the PM task, fewer should be available for ongoing task decisions, which should result in slower evidence accumulation on lexical decision trials. Alternatively, if cue detection occurs because participants endogenously implement delays in responding to reduce response competition (as suggested by the delay theory), this implementation should result in higher thresholds for responding (i.e., boundary separation). Across several studies the authors found that in nonfocal conditions ongoing task cost was largely associated with increased decision boundaries with no changes in drift rate. Strickland, Heathcote, Remington, and Loft (2017) replicated these findings and also found that focal costs, although relatively minimal, could also be accounted for by increases in boundary separation.[1] These findings were taken as evidence against the prevailing view the ongoing and PM tasks compete for shared resources[2] and instead suggest that cue detection is a result of increased response thresholds that allow more time for PM information to accumulate and complete for response selection.

It is important to note that the aforementioned application of evidence accumulation models, and consequently the primary evidence in favor of delay theory, has only relied upon examination of noncue trials in PM tasks (Heathcote et al., 2015; Strickland et al., 2017). That is, evidence accumulation models are typically applied to standard ongoing task trials to account for nonfocal cost effects. Based on the prevailing view that slowing underlies nonfocal cue detection, it was inferred that increased response thresholds was the primary contributor to PM performance on cue trials that were not actually modeled in this approach. A more recent model variant, the prospective memory decision control (PMDC) model, explicitly accounts for both ongoing task and PM trial responding (Strickland, Loft, Remington, & Heathcote, 2018; see also Boag, Strickland, Heathcote, Neal, & Loft, 2019; Boag, Strickland, Loft, & Heathcote, 2019; Strickland et al., 2019). The PMDC model suggests that ongoing task information (e.g., word and nonword accumulators in a lexical-decision task) and PM information (i.e., PM accumulator) accumulates in parallel and competes for response selection. As such, it retains the core assumption of delay theory in regard to the role of response thresholds in allowing more time for PM information to accumulate. Critically, the model additionally suggests that to increase the likelihood that PM response selection occurs, on PM trials there may be excitation of the PM accumulator (e.g., speeding accumulation of PM information) or inhibition of ongoing task accumulators (e.g., slowing accumulation of non-PM information).

Using the PMDC framework, Strickland et al. (2018) fit the linear ballistic accumulator model to PM data that included a control, focal PM, and nonfocal PM block. Consistent with prior research, it was found that slowing in the nonfocal block was associated with threshold increases. In addition, on both focal and nonfocal PM cue trials there was greater excitation of the PM accumulator and inhibition of ongoing task (word and nonword) accumulators. Somewhat surprisingly, however, threshold changes were not predictive of PM performance. Rather, inhibition of ongoing task accumulators on cue trials was predictive of performance. These findings suggest that multiple processes (delay and inhibition) may operate during PM tasks (Boag, Strickland, Heathcote, et al., 2019; Boag, Strickland, Loft, & Heathcote, 2019; Strickland et al., 2019), but call into question whether a delay mechanism is truly functional for PM performance (see also Anderson, Rummel, & McDaniel, 2018). Because the processes underlying performance on the experimental task using the tone procedure (Loft & Remington, 2013) do not necessarily map onto the parameter estimates derived from modeling work (Strickland et al., 2018), however, more experimental work is needed to test whether a delay mechanism is beneficial for PM performance.

Current Study

In our view, the Loft and Remington (2013) results are some of the most interesting and theoretically important PM findings in recent literature. This work provided the foundation for the development of new theory and modeling of PM, including the delay theory (Heathcote et al., 2015) and the PMDC model (Strickland et al., 2018). These findings are pushing the boundaries of PM theorizing and have caused PM researchers to think more deeply about traditional ideas of PM processing. Given the significance of such findings, we thought it was prudent to revisit the theoretical mechanisms originally described by Loft and Remington. To our knowledge, the original delay procedure has only been examined in one other study. Loft, Bowden, Ball, and Brewer (2014) examined PM performance in HIV+ individuals using the syllable nonfocal cue task described earlier. It was found that HIV+ individuals showed reduced focality effects with increased delays. This suggests that the reduced focality effect with increased delay is replicable, as a total of four different experiments across two independent samples have shown this finding (Loft & Remington, 2013; Loft et al., 2014). The purpose of the current study was to provide additional support for this idea and to further test claims that a delay mechanism is a viable mechanism contributing to PM performance.

Notably, although Loft et al. (2014) found that delay reduced the focality effect, in contrast to Loft and Remington (2013) the effect was never completely eliminated even at the longest delays (1,600 ms). Loft et al. (2014) posited that with increased delays (e.g., 2,500 ms) presumably this effect would be eliminated. Although not an explicit assumption by Loft and Remington (2013), or necessarily of the delay theory more broadly, we believe that the complete elimination of the focality effect is a critical test of the viability of a delay mechanism in explaining PM performance. Delay theory is a single process theory that suggests that response thresholds are the sole contributor to cue detection (Heathcote et al., 2015; Strickland et al., 2017). This means that if enough time

[1] It should be noted, however, at least two other studies have found that manipulations thought to influence the degree of monitoring enacted (e.g., importance of intention, cue frequency, and cue focality) produced changes in nondecision time (Ball & Aschenbrenner, 2018; Horn & Bayen, 2015). These findings have been interpreted as reflecting that PM-specific monitoring processes (i.e., checking for PM cues) may occur either prior to or following the ongoing task decision. This is entirely consistent with traditional theories of PM, and intuitively makes sense. However, Strickland et al. (2017) have argued that these findings simply reflect an artifact of the type of model that was used (i.e., diffusion model rather than linear ballistic accumulator model).

[2] Recent work with more challenging ongoing tasks (e.g., air traffic control simulation) does show evidence of capacity sharing (Boag, Strickland, Heathcote, et al., 2019; Boag, Strickland, Loft, & Heathcote, 2019; Strickland et al., 2019). It is likely that these more complex tasks are closer to human capacity limits than in the traditional lexical decision task used for studying PM, and therefore provide conditions where the consequences of capacity sharing can manifest.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

is given for PM information to accumulate, nonfocal cue detection should always approximate focal cue detection and both should reach ceiling. Thus, in addition to consistently demonstrating a reduced focality effect with increased delays, complete elimination of the effect at longer intervals would provide strong support for the delay theory. If instead the effect is never eliminated (e.g., there is a point of diminishing return), this would suggest that some other mechanism also contributes to performance (e.g., reactive inhibition, Strickland et al., 2018).

Across five experiments, we tested the generalizability of the original findings of Loft and Remington (2013). Experiments 1–4 examined boundary conditions of the delay theory, whereas Experiments 5 more directly tested contrasting theoretical predictions from monitoring theories of PM (e.g., multiprocess framework) and threshold theories of PM (i.e., delay theory). Experiment 1 was a close replication of the delay procedure used in the original paper, whereas Experiment 2 used a slight variant of the tone procedure while maintaining the other critical aspects of the original procedure. Using the modified tone procedure Experiments 3 and 4 implemented a different set of PM cues and increased the length of the delays (up to 2,500 ms), respectively. Based on the original findings, we anticipated that the focality effect would be reduced with increased delays and completely eliminated by 1,600 ms. In Experiment 5 we adopted an individual differences approach to further test the theoretical mechanisms underlying the delay effect (and PM more generally). Monitoring theories of PM predict that individuals with poor attention abilities should not show a benefit to PM with increased delays, whereas the delayed responding theories predict that they should benefit most from delays. Thus, Experiment 5 assessed attention ability and PM performance using a large-scale individual differences study in a group of younger adults.

Experiment 1

Experiment 1 was a close replication of the original procedure used in Experiment 3 by Loft and Remington (2013), with the exception that we did not include a no-intention control block since we were not particularly interested in ongoing task costs. Participants performed an ongoing lexical-decision task in both focal and nonfocal blocks. The tone was presented 0 ms (i.e., immediately), 600 ms, 1,000 ms, and 1,600 ms following stimulus onset and was randomly selected on each trial. In the focal block participants were given an intention to make a special response to a specific word (e.g., operator), whereas in the nonfocal block participants were to respond any time a specific syllable appeared (e.g., fer) within one of the ongoing task stimuli. There were a total of eight cues presented (two for each delay) in each block. As is typical in the PM literature, in the focal block a single cue (e.g., operator) repeated across all eight trials, whereas in the nonfocal block eight unique cues were presented (e.g., offer, feral, ferret, etc.). Based on the original findings, we anticipated that the focality effect would be reduced by 600 ms and eliminated by 1,600 ms.

We also included an additional between-subjects condition using a nearly identical procedure. The only difference was that in the nonfocal block, one of the eight syllable cue word (e.g., offer) was selected for each participant that repeated across all eight trials. That is, participants were instructed to look for a particular syllable (e.g., fer) but only one PM cue (e.g., offer) was repeated throughout the task (e.g., feral, ferret, etc., were not presented). This manipulation makes the nonfocal block more similar to the focal block by controlling for stimulus exposure effects (Hicks, Franks, & Spitler, 2017). Although not considered in prior modeling work, presumably repeated exposure to the same focal (but not nonfocal) cue may speed evidence accumulation rates or reactive inhibition on cue trials with experience, producing a selective benefit for focal cue detection. Including the repeated nonfocal cue condition ensures that any influence on performance can be directly attributed to the delay manipulation rather than simple stimulus repetition effects. To foreshadow, however, this manipulation had no influence on performance.

Method

All research reported herein was conducted using appropriate ethical guidelines and was approved by the Institutional Review Board at each university where the data was collected (Arizona State University, Washington University in Saint Louis, or the University of Texas at Arlington).

Participants and design. Seventy-seven undergraduates from the University of Texas at Arlington received course credit for participation. Participants were randomly assigned to the unique ($n = 40$) and repeated nonfocal ($n = 37$) cues condition. Although the total sample size was chosen to approximate the sample size of Loft and Remington (2013), it should be noted that the replication condition (i.e., unique condition) actually has fewer participants than the original reported study (which consisted of 72 participants). However, based on the effect size ($\eta_p^2 = .07$) from Loft and Remington (Experiment 3) only 32 participants are needed to detect the interaction effect with .80 power. This power analysis was the basis for our sample size in Experiments 1–4. All participants completed the focal and nonfocal cue blocks (order counterbalanced).

Materials. The materials and stimuli were created based on the methods provided in Experiment 3 of Loft and Remington (2013). Unlike Loft and Remington (2013), this experiment did not include control blocks. All stimuli and response collection was done through E-prime (Schneider, Eschman, & Zuccolotto, 2002).

For the lexical-decision task stimuli, we randomly selected 428 medium frequency words ($M = 43$ occurrences per million, $SD = 30$) from the English Lexicon Project database (Balota et al., 2007) and replaced one vowel from each word (e.g., chart, regard) with another to produce 428 pronounceable nonwords (e.g., chirt, ragard). The computer program randomly selected for each participant one of the trial types (e.g., chart, ragard) to be presented, meaning that the other trial type (e.g., chirt, regard) would not be presented, with the stipulation that across the entire experiment half of the items were words and half were nonwords.

Each block of the lexical decision task (LDT) consisted of 210 trials. Presentation order of words and nonwords within each list was random, with the exception of PM target trials which were presented on Trials 30, 55, 80, 105, 130, 155, 180, and 205. Three target syllables (*fer, tor,* and *ver*) were selected to create three lists of PM cues. As in Loft and Remington (2013), in Set A the target syllable was *fer* (feral, ferry, offering, prefer, refer, referee, suffer, transfer), in Set B the target syllable was *tor* (victory, actor, investor, editor, historic, factory, operator, monitor), and in Set C the target syllable was *ver* (coverage, verdict, poverty, deliver, clever, forever, version, recover). One set was assigned to the focal block (e.g., Set A) and a second set was assigned to the nonfocal block (e.g., Set B), with a total of 6 possible combinations that was counterbalanced across participants (i.e., A-B, A-C, B-A, B-C, C-A, or C-B). In the focal block, participants were instructed to press the "7" key whenever a specific word appeared (e.g., feral), whereas in the nonfocal block participants were instructed to press the "7" key whenever a specific syllable appeared within one of the words (e.g., tor). The PM response was to be made *instead* of the word/nonword response and after the tone had played. In the focal block, the computer randomly selected one of the eight cues from the category (e.g., feral) to be presented on all eight PM cue trials. In the nonfocal unique condition (replicating Loft & Remington, 2013), each of the eight cues from the category (e.g., victory, actor, investor, etc.) was randomly assigned to one of the eight PM cue trials. In the nonfocal repeated condition, the computer randomly selected one of the eight cues from the category (e.g., victory) to be presented on all eight PM cue trials. Thus, in both nonfocal blocks participants were instructed to look out for a syllable (e.g., tor), but in the unique condition the *tor* syllable was in a different word each time whereas in the repeated condition the *tor* syllable was in the same word each time.

Four delays (0 ms, 600 ms, 1,000 ms, and 1,600 ms) were randomly assigned to nontarget trials, and each tone delay was presented equally often. The tone delay presented on target trials was systematically manipulated. For the first four presentations of targets, each tone delay was presented once in a random order. In the next four presentations of targets, each tone delay was presented once again in a random order.

Procedure. Participants were presented with either a word or nonword and asked to accurately determine into which category it belonged. The decision was rendered by the participant by either pressing the "F" key for word and the "J" key for a nonword. Each trial began with a fixation cross (i.e., "+") at the center of the screen for 250 ms. The fixation was then replaced with either a word or nonword and after a 0, 600, 1,000, or 1,600 ms a 500-Hz tone was then played. Two of the four delays were randomly select to occur 52 times per block, whereas the other two were randomly selected to occur 53 times. These extra trials had no bearing on PM performance. The word/nonword remained on the screen until the participant responded. Participants were instructed to respond as quickly and accurately as possible, but only after the tone had been played.

Participants completed 24 practice trials. After the practice trials, participants completed three blocks of experiment trials. After the practice trials the participant was given the instructions for the PM blocks. During the PM blocks they were instructed to continue making decisions about whether the word presented was a word or a nonword but if they encountered a word that fit the criteria for either a focal or nonfocal PM cue, the participant should hit a secondary key (i.e., "7" key). Under the focal conditions they identified when the specific cue word was shown (e.g., feral). During the nonfocal conditions they identified when the word shown contains a specific syllable (e.g., "tor"). As with standard LDT trials, participants were instructed to withhold their PM response until after the tone had played. After receiving the PM instructions and prior to the focal or nonfocal blocks, the participant did a brief distractor task (Shipley's Vocabulary Test during each). Specifically, the participant was shown a word and four response options. The participant was tasked with selecting the option that was a synonym for the word shown. The participant completed 20 distractor items (40 items total) before each PM block. The distractor was self-paced but took approximately 2.5 min to complete.

Results

Cue detection, ongoing task accuracy, and ongoing task response times (RTs) were submitted separately to a 2 (cue type: focal vs. nonfocal; within-subjects) $\times$ 4 (delay: 0 vs. 600 vs. 1,000 vs. 1,600 ms; within-subjects) $\times$ 2 (exposure: unique vs. repeated; between-subjects) mixed-factorial analysis of variance (ANOVA). Results can be found in Figure 1. Unless otherwise noted, the alpha level was set at .05 for all analyses.

Cue detection. *PM accuracy* was defined as the proportion of PM cue trials receiving a "7" response following the tone. Late PM responses (within two trials of the PM cue) were rare (M = .053, SE = .008) and were counted as correct (Loft & Remington, 2013).

Cue detection was higher for focal than nonfocal cues [cue type: $F(1, 75) = 23.28$, $p < .001$, $\eta_p^2 = .237$]. Cue detection did not increase with delay [delay: $F(3, 225) = 1.24$, $p = .296$, $\eta_p^2 = .016$]. Critically, the focality effect was not reduced with delays [Cue Type $\times$ Delay: $F(3, 225) < 1$]. Although cue detection was better overall when nonfocal cues were repeated [exposure: $F(1, 75) = 4.17$, $p = .045$, $\eta_p^2 = .055$], this exposure effect did not interact

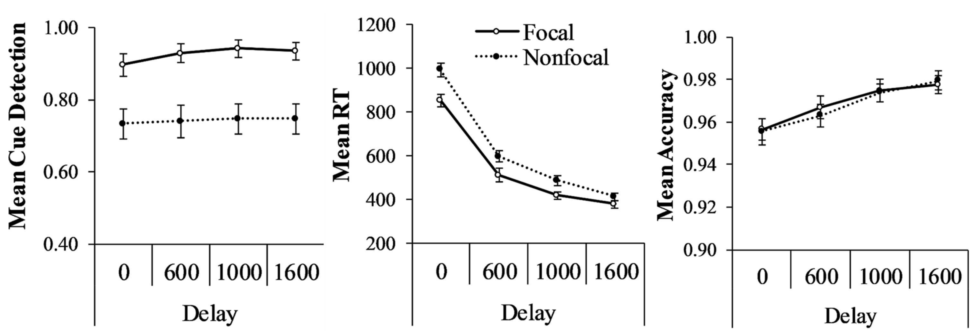

Figure 1. Mean cue detection (left panel), ongoing task reaction times (RTs; middle panel), and ongoing task accuracy (right panel) as a function of cue type (focal vs. nonfocal) and delay in Experiment 1. Error bars reflect standard errors.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

with cue type or delay [Exposure × Cue Type: $F < 1$; Exposure × Delay: $F(3, 225) = 1.24$, $p = .296$, $\eta_p^2 = .016$; Exposure × Cue Type × Delay: $F < 1$]. Planned contrasts revealed that focal cue detection was higher than nonfocal cue detection at all delays (all ps $< .003$).

Ongoing task performance. Following Loft and Remington (2013), ongoing task performance was assessed for word trials only. PM cue trials and the four trials following cues were excluded. RT analyses were conducted on correct trials only, and for each delay RTs to word trials greater than three standard deviations from a participant's grand mean for that delay were excluded. Ongoing task analyses were performed on the reduced sample used in the cue detection analyses.

RT. RT was faster in the focal block [cue type: $F(1, 75) = 24.83$, $p < .001$, $\eta_p^2 = .249$] and decreased with delay [delay: $F(3, 225) = 459.17$, $p < .001$, $\eta_p^2 = .860$]. Critically, the focal RT advantage was reduced with delay [Cue Type × Delay: $F(3, 225) = 8.40$, $p < .001$, $\eta_p^2 = .101$]. Repeating nonfocal cues had no influence on performance [exposure: $F < 1$; Exposure × Cue Type: $F < 1$; Exposure × Delay: $F < 1$; Exposure × Cue Type × Delay: $F(3, 225) = 1.95$, $p = .123$, $\eta_p^2 = .025$]. The Cue × Delay interaction reflects that although focal RTs were faster than nonfocal RTs at all delays (all ps $< .013$), the magnitude of the effect decreased across delays.

Accuracy. Accuracy did not differ across blocks [cue type: $F < 1$] but did increase with delay [delay: $F(3, 225) = 8.72$, $p < .001$, $\eta_p^2 = .104$]. Critically, the accuracy focality effect was not reduced with delay [Cue Type × Delay: $F < 1$]. Although repeating nonfocal cues did not influence overall accuracy [exposure: $F(1, 75) = 1.65$, $p = .203$, $\eta_p^2 = .022$], cue repetition interacted with cue type [Exposure × Cue Type: $F(1, 75) = 5.74$, $p = .019$, $\eta_p^2 = .071$]. This interaction reflects that accuracy was higher in the repeated than the unique condition during the focal block ($p = .025$) but did not differ in the nonfocal block ($p = .757$). Repeating nonfocal cues had no other influence on performance [Exposure × Delay: $F < 1$; Exposure × Cue Type × Delay: $F < 1$].

Discussion

The results from Experiment 1 are relatively straightforward. As is typically the case in PM, focal cue detection was considerably higher than nonfocal cue detection. Somewhat surprisingly, however, the difference between focal and nonfocal cue detection (i.e., the focality effect) was not reduced with increased delays. As mentioned previously, Loft et al. (2014) replicated the original findings of Loft and Remington (2013) by showing a significant interaction effect such that delay reduced the focality effect (although the effect was never completely eliminated). Thus, it is not entirely clear why we did not find a similar reduction, especially considering that delays did reduce ongoing task RT differences typically seen between focal and nonfocal conditions. The results do rule out any stimulus exposure account for the current or previous findings, as repeating nonfocal cues in a similar manner to how focal cues had little impact on performance. Experiment 2 was designed to replicate the current findings using a slightly different tone procedure.

Experiment 2

Experiment 2 was a replication of Experiment 1 but using a modified tone procedure. In Experiment 2, the tone onset was at the same time as the lexical stimulus presentation for all delays but was played for variable durations (600 ms, 1,000 ms, and 1,600 ms). That is, for all delay trials the stimulus (e.g., word) and tone onset at the same time and the tone continued playing for some duration (e.g., 1,000 ms). Participants were instructed to withhold responding until after the tone had finished playing. On no-tone trials (0 ms), participants were instructed they could respond immediately. In this regard, the critical delay aspect was maintained without producing selective interference on the 0 ms condition. That is, the tone in Experiment 1 influenced initial processing of stimulus information only during the 0 ms trial because the tone and stimulus onset at the same time, whereas on all other trials initial stimulus processing was left unaffected. With this new procedure, stimulus processing is equally affected across all delays trials because the tone is continuously played. Based on Experiment 3 of Loft and Remington (2013), we anticipated that the focality effect would be reduced by 600 ms and eliminated by 1,600 ms. However, given the findings of Experiment 1 we also thought it was possible to find no influence of delay on the focality effect.

Method

Participants and design. Eighty-two undergraduates from the University of Texas at Arlington received course credit for participation. Participants were randomly assigned to the unique ($n = 43$) and repeated nonfocal ($N = 40$) cues condition. Although the total sample size was chosen to approximate the sample size of Loft and Remington (2013), as with Experiment 1 it should be noted that the replication condition (i.e., unique condition) actually has fewer participants than the original reported study (which consisted of 72 participants). All participants completed the focal and nonfocal cue blocks (order counterbalanced).

Materials and procedure. The materials and procedure were identical to Experiment 1 with one modification to the tone procedure. As described previously, in Experiment 2 the tone onset at the same time as the lexical stimulus for all delays, but was played for variable durations (0 ms, 600 ms, 1,000 ms, and 1,600 ms). Participants were instructed to withhold responding until after the tone had finished playing on delay trials (600, 1,000, and 1,600 ms), or that they could respond immediately on no-tone trials (0 ms). More explicitly, participants were instructed that if they did not hear a tone when the word or nonword appeared, that meant a tone would not be played on that trial and they could respond immediately. Otherwise, they should wait to respond until the tone had finished playing.

Results

Late PM responses (within two trials, $M = .025$, $SE = .006$) were rare and counted as correct. Cue detection, ongoing task accuracy, and ongoing task RT were submitted separately to a 2 (cue type: focal vs. nonfocal; within-subjects) × 4 (delay: 0 vs. 600 vs. 1,000 vs. 1,600 ms; within-subjects) × 2 (exposure: unique vs. repeated; between-subjects) mixed-factorial ANOVA. Results can be found in Figure 2.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

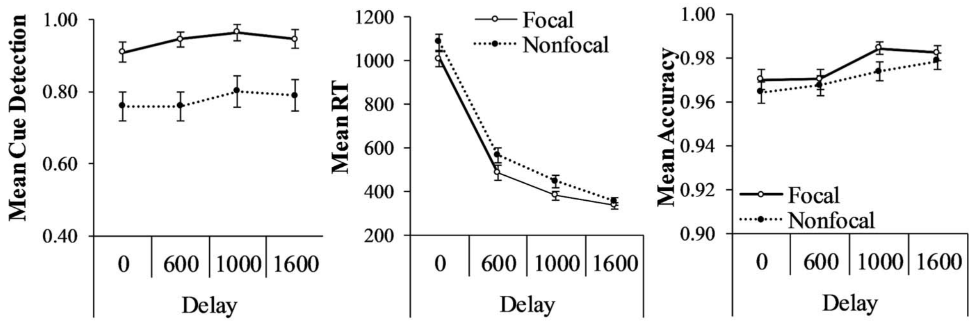

Figure 2. Mean cue detection (left panel), ongoing task reaction times (RTs; middle panel), and ongoing task accuracy (right panel) as a function of cue type (focal vs. nonfocal) and delay in Experiment 2. Error bars reflect standard errors.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Cue detection. Cue detection was higher for focal than nonfocal cues [cue type: $F(1, 81) = 22.11$, $p < .001$, $\eta_p^2 = .214$]. However, there were no other significant effects. Cue detection did not increase with delay [delay: $F(3, 243) = 1.89$, $p = .132$, $\eta_p^2 = .023$], the focality effect was not reduced with delays [Cue Type × Delay: $F < 1$], and repeating nonfocal cues had no influence on performance [exposure: $F < 1$; Exposure × Cue Type: $F < 1$; Exposure × Delay: $F < 1$; Exposure × Cue Type × Delay: $F < 1$]. Planned contrasts revealed that focal cue detection was higher than nonfocal cue detection at all delays (all *p*s < .002).

Ongoing task performance.

RT. RT was faster in the focal block [cue type: $F(1, 81) = 14.51$, $p < .001$, $\eta_p^2 = .152$] and decreased with delay [delay: $F(3, 243) = 546.69$, $p < .001$, $\eta_p^2 = .871$]. Critically, the focal RT advantage was reduced with delay [cue type x delay: $F(3, 243) = 3.30$, $p = .021$, $\eta_p^2 = .038$]. Repeating nonfocal cues had no influence on performance [exposure: $F < 1$; Exposure × Cue Type: $F(1, 81) = 1.79$, $p = .184$, $\eta_p^2 = .022$; Exposure × Delay: $F < 1$; Exposure × Cue Type × Delay: $F < 1$]. The Cue × Delay interaction reflects that focal RTs were faster than nonfocal RTs with delay intervals of 0 ms, 600 ms, and 1,000 ms (all *p*s < .002), but not with a delay of 1,600 ms ($p = .074$).

Accuracy. Accuracy was marginally higher in the focal block [cue type: $F(1, 81) = 3.42$, $p = .068$, $\eta_p^2 = .041$]. Accuracy also increased with delay [delay: $F(3, 243) = 6.40$, $p < .001$, $\eta_p^2 = .073$]. However, there were no other significant effects. The accuracy focality effect was not reduced with delays [Cue Type × Delay: $F < 1$], and repeating nonfocal cues had no influence on performance [exposure: < 1; Exposure × Cue Type: $F(1, 81) = 2.23$, $p = .140$, $\eta_p^2 = .027$; Exposure × Delay: $F < 1$; Exposure × Cue Type × Delay: $F(1, 81) = 1.05$, $p = .371$, $\eta_p^2 = .013$].

Discussion

The results from Experiment 2 using the modified tone procedure replicate the primary findings of Experiment 1. Focal cue detection was considerably higher than nonfocal cue detection, but somewhat surprisingly, was not reduced with delays. Although delay did not influence cue detection, it did reduce ongoing task RT differences between focal and nonfocal conditions. Lastly, the cue repetition manipulation had no influence on performance. Given the two failed replications of the original findings, Experiment 3 used a slightly different set of cues to assess the influence of delay on performance.

Experiment 3

Loft and Remington (2013) found that the focality effect was eliminated by 600 ms with a categorical cue, but it took 1,600 ms with a syllable cue. This is consistent with previous research indicating that syllable information is more difficult than words (Anderson et al., 2017, 2019) and suggests that the exact properties of the PM cues have some influence on the delayed focality effect. Experiment 3 was designed to replicate the findings of Experiment 2 using the modified tone procedure but with a slightly different set of cues. In the focal block the cue was always the word *sister*, whereas in the nonfocal block the cue was always any word starting with the letter "S". These cues were selected because previous research has indicated that monitoring for words and first letter information is of comparable difficulty (Scullin, McDaniel, Shelton, & Lee, 2010). Despite the comparable difficulty, the ongoing lexical-decision task does not orient attention to the first letter on each trial and should result in worse cue detection. Based on Experiment 3 of Loft and Remington (2013), we anticipated that the focality effect would be reduced by 600 ms and eliminated by 1,600 ms. However, given the findings of Experiment 1 and 2 we also thought it was possible to find no influence of delay on the focality effect. We did not include the between-subjects manipulation from Experiments 1 and 2 because stimulus exposure had no influence on performance.

Method

Participants and design. Forty undergraduates from Washington University in Saint Louis received course credit for participation. This sample size was chosen to be consistent with the unique cue condition of the previous experiments. All participants completed the focal and nonfocal cue blocks (order counterbalanced).

Procedure. Experiment 3 used the same materials and procedure from Experiment 2. The focal and nonfocal cues were changed. In the focal block the PM cue was the word *sister*. In the nonfocal block, the PM cue was any word starting with "S" (spleen, smoke, sugar, sand, silver, stereo, shoulder, spring).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Results

Late PM responses (within two trials, M = .067, SE = .015) were rare and counted as correct. Although rare, there were also a few instances where participants erroneously made a lexical decision response on PM trials before the tone ended (M = .003, SE = .002). Following Loft and Remington (2013), we excluded these trials because this reflects cases where the full delay was not utilized. Cue detection, ongoing task accuracy, and ongoing task RTs were submitted separately to a 2 (cue type: focal vs. nonfocal; within-subjects) × 4 (delay: 0 vs. 600 vs. 1,000 vs. 1,600 ms; within-subjects) ANOVA. Results can be found in Figure 3.

Cue detection. Cue detection was higher for focal cues [cue type: $F(1, 39) = 35.35$, $p < .001$, $\eta_p^2 = .475$] and increased with delays [delay: $F(3, 117) = 4.06$, $p = .009$, $\eta_p^2 = .094$]. However, the focality effect was not reduced with delays [Cue Type × Delay: $F(3, 117) = 1.66$, $p = .180$, $\eta_p^2 = .041$]. Planned contrasts revealed that focal cue detection was higher than nonfocal cue detection at all delays (all ps < .006).

Ongoing task performance.

RT. RT decreased with delay [delay: $F(3, 117) = 412.09$, $p < .001$, $\eta_p^2 = .914$]. However, there were no other significant effects. RTs were not faster in the focal block [cue type: $F < 1$] and this did not change across delays [Cue Type × Delay: $F < 1$].

Accuracy. Accuracy increased with delay [delay: $F(3, 117) = 10.92$, $p < .001$, $\eta_p^2 = .219$]. However, there were no other significant effects. Accuracy was not better in the focal block [cue type: $F(1, 39) = 1.93$, $p = .173$, $\eta_p^2 = .047$] and this did not change across delays [Cue Type × Delay: $F < 1$].

Discussion

The results from Experiment 3 are consistent with the previous experiments in that there was no reduction in the focality effect with increased delays. In contrast to Experiment 2, delay did not influence ongoing performance differentially across blocks. That is, responding was equally as fast and accurate in the focal and nonfocal blocks. We suspect this in part reflects the choice of cues, which were specifically selected because previous research showed that monitoring across words and first letters is of comparable difficulty (Scullin et al., 2010). Although the reduction in the focality effect was not significant, the results trended in the right direction. It is possible that the delay used was simply not long enough to show a benefit to performance. Thus, Experiment 4 replicated the current procedure but increased the delay intervals to 2,500 ms.

Experiment 4

Because 1,600 ms may have not been enough time for PM response selection to occur, Experiment 4 replicated the procedure from Experiment 3 but extended the delays up to 2,500 ms. The delay intervals in Experiment 4 were 0 ms (no tone), 1,000 ms, 1,500 ms, 2,000 ms, and 2,500 ms. Based on delay theory predictions, we anticipated that the focality effect should be reduced by 1,500 ms and eliminated by the later delay intervals (2,000 or 2,500 ms).

Method

Participants and design. Forty-one undergraduates from Washington University in Saint Louis received course credit for participation. This sample size was selected based on the previous experiment. All participants completed the focal and nonfocal cue blocks (order counterbalanced).

Procedure. Experiment 4 used the same procedure and word/nonword list as Experiment 3. However, the length of the delay was changed. In Experiments 2 & 3 the lengths of the delays were 0, 600, 1,000, or 1,600 ms. In this experiment, the lengths of the delays were 0, 1,000, 1,500, 2,000, 2,500 ms. In addition, each block now consisted of 310 trials due to the addition of a fifth delay interval. In the focal block the PM cue was the word *sister*. In the nonfocal block, the PM cue was any word starting with "S" (spleen, smoke, sugar, sand, silver, stereo, shoulder, spring, skate, and speaker).

Results

Late PM responses (within two trials, M = .054, SE = .020) were rare and counted as correct. Although rare, there were also a few instances where participants erroneously made a lexical decision response on PM trials before the tone ended (M = .001, SE = .001). Following Loft and Remington (2013), we excluded these trials because this reflects cases where the full delay was not utilized. Cue detection, ongoing task accuracy, and ongoing task response times (RTs) were submitted separately to a 2 (cue type:

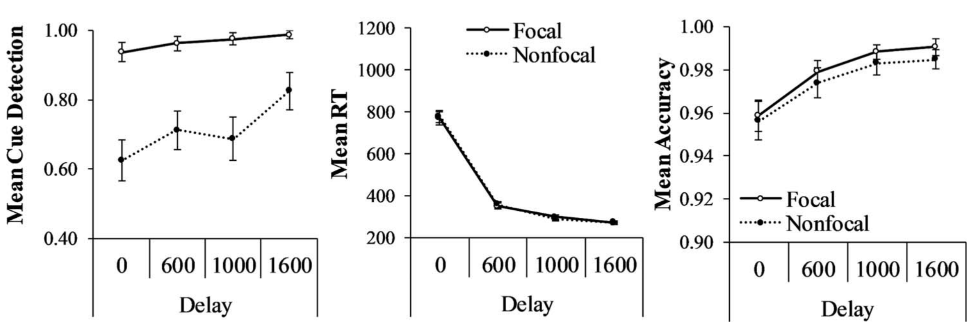

Figure 3. Mean cue detection (left panel), ongoing task reaction times (RTs; middle panel), and ongoing task accuracy (right panel) as a function of cue type (focal vs. nonfocal) and delay in Experiment 3. Error bars reflect standard errors.

focal vs. nonfocal; within-subjects) × 5 (delay: 0 vs. 1,000 vs. 1,500 vs. 2,000 vs. 2,500 ms; within-subjects) ANOVA. Results can be found in Figure 4.

Cue detection. Cue detection was higher for focal cues [cue type: $F(1, 40) = 45.38$, $p < .001$, $\eta_p^2 = .532$]. However, performance did not increase with delays [delay: $F(4, 160) = 1.52$, $p = .198$, $\eta_p^2 = .037$] and the focality effect was not reduced with delays [Cue Type × Delay: $F(4, 160) = 1.99$, $p = .099$, $\eta_p^2 = .047$]. Planned contrasts revealed that focal cue detection was higher than nonfocal cue detection at all delays (all ps $< .002$).

Ongoing task performance.

RT. RT decreased with delay [delay: $F(4, 160) = 269.53$, $p < .001$, $\eta_p^2 = .871$]. However, RTs were not faster in the focal block [cue type: $F(1, 40) = 1.15$, $p = .289$, $\eta_p^2 = .028$]. and this did not change across delays [Cue Type × Delay: $F < 1$].

Accuracy. Accuracy increased with delay [delay: $F(4, 160) = 8.74$, $p < .001$, $\eta_p^2 = .179$]. However, there were no other significant effects. Accuracy was not better in the focal block [cue type: $F < 1$] and this did not change across delays [Cue Type × Delay: $F(4, 160) = 1.12$, $p = .349$, $\eta_p^2 = .027$].

Discussion

The results of Experiment 4 are consistent with those of Experiment 3. Examining either the full range of delays or the subset of delays most comparable to previous experiments (0, 1,000, 1,500 ms), there was no elimination of the focality effect with increased delay. Moreover, delay did not influence ongoing performance differentially across blocks. Critically, Experiment 4 demonstrated that extending the delay intervals longer (2,000–2,500 ms) showed no additional benefit to nonfocal PM performance. In fact, there was even a tendency to do worse with longer delays. This finding is particularly difficult to reconcile from a delay account of PM, as there should have been ample time for response selection to occur.

Experiment 5

The previous experiments were designed to replicate and extend the original findings reported by Loft and Remington (2013). Unfortunately, and somewhat surprisingly, we were largely unsuccessful at replicating those findings. Experiment 5 takes a slightly different approach to more directly test predictions from different theories of PM. To do this, we assessed individual differences in attention ability (working memory, proactive control, inhibition, and task switching) in a large group of college-aged individuals. Considerable research suggests that individuals with low attention ability typically show deficits in nonfocal tasks. For example, Brewer et al. (2010) found that high working memory capacity (WMC) individuals detected more nonfocal cues but similar numbers of focal cues as low WMC individuals. It is generally thought that low ability participants are less able to appropriately sustain monitoring throughout the nonfocal PM task while simultaneously performing a demanding ongoing task. This finding—substantial ability differences in nonfocal cue detection but negligible effects in focal cue detection—has often been used to support the theoretical proposal that cue detection can occur via two distinct processes: monitoring and spontaneous retrieval. Alternatively, it has been suggested by proponents of the delay theory that "executive capacity may be used to adjust thresholds to meet PM task demands" and that low ability participants may "fail to increase their thresholds in prospective memory blocks" (Strickland et al., 2017, p. 9). Thus, the finding that low ability participants do worse on the more demanding nonfocal task may reflect that these individuals do not set the appropriate response thresholds to allow enough time for PM information to accumulate.

Based on these views, there are two alternative predictions that can be made. From a monitoring theory perspective, attention ability should *not* moderate the effect of delay on focality because delaying responding should not increase the likelihood of maintaining the PM intention. That is, high ability participants should show better nonfocal performance across all delays. The rationale here is that monitoring is independent of delay durations—if one is not good at sustaining monitoring across trials, it should not matter whether there is a 0 ms delay or 1,600 ms delay. Based on delay theory, attention ability *should* moderate the delayed focality effect. The rationale here is that low ability participants are either less likely or less able to endogenously set appropriately optimal response thresholds (Strickland et al., 2017). The tone procedure circumvents these limitations by exogenously setting response thresholds higher than what a low ability participant would endogenously set on their own (i.e., during the 0 ms delay). This increased threshold allows more time for PM information to accumulate, meaning that low ability participants should show considerable improvements across delays. In contrast, because high ability participants are already setting appropriate response thresholds without tones (i.e., during the 0-ms delay), forcing them to

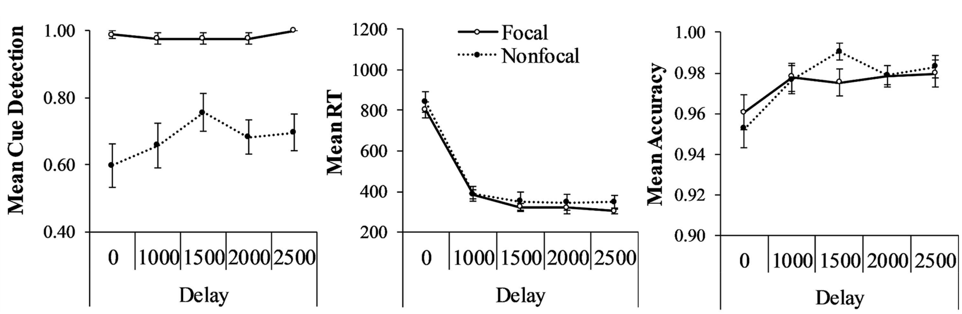

Figure 4. Mean cue detection (left panel), ongoing task reaction times (RTs; middle panel), and ongoing task accuracy (right panel) as a function of cue type (focal vs. nonfocal) and delay in Experiment 4. Error bars reflect standard errors.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

withhold responding until a tone has played should not produce as large of a benefit to performance.

To test these alternative predictions, we used a similar procedure as Experiment 3 but only included delays of 0 ms (no tone), 1,000 ms, and 1,600 ms. We reduced the total number of different durations to three so we could increase the number of cues presented at each duration (six instead of two) per block to get more stable estimates of PM without making the task exceedingly long. In addition to more PM measures than previous experiments, we included a large sample size (N = 249). Although the overall sample size collapsed across exposure condition of Experiments 1 and 2 was comparable to Loft and Remington (2013), one potential criticism of the Experiments 3 and 4 is that they were underpowered to detect significant effects (both experiments trended in the direction originally reported by Loft and colleagues). This should not be a concern in the current experiment. The large sample size was also chosen because in addition to the PM task, participants performed a battery of cognitive tasks to assess different attention abilities. All participants completed a verbal and a visuospatial version of tasks used to assess WMC (complex span task), proactive control (AX-continuous performance task [CPT]), inhibition (sustained attention to response task), and task-switching (switch task). We selected these attention constructs because previous research has shown that each is important for PM. It has been suggested that WMC is involved in maintenance of the PM intention (Brewer et al., 2010), proactive control and inhibition are involved in inhibiting prepotent ongoing task responses to check for PM cues (Ball & Brewer, 2018; Zuber, Kliegel, & Ihle, 2016), and task-switching is involved in switching from the ongoing task to the PM task (Schnitzspahn et al., 2013; Zuber et al., 2016). Moreover, working memory, proactive control, and inhibition have all been directly tied to modeling parameters (Strickland et al., 2017, 2018). Although we generally expected each of these abilities to be associated with PM (particularly for nonfocal PM), we did not have any a priori predictions that one ability (e.g., WMC) would necessarily influence the delayed focality effect differently than another (e.g., task switching). Rather, these abilities were chosen to broadly cover different aspects of cognition associated with PM and to hopefully provide converging evidence of the true effects across multiple measures. Any differences that may arise across abilities, however, could be informative in constraining theory testing.

Method

Participants and design. Two hundred and ninety-two undergraduate participants from Arizona State University received course credit for participation. The increased sample size for this experiment was to ensure the proper power to evaluate individual differences in attention control ability. Data were excluded from analyses from 18 participants with missing data on one or more of the tasks and from 25 participants that were identified as multivariate outliers using Mahalonbis distance estimates between the two tasks for each construct (nine based on proactive control, nine based on task switching measures, and seven based on inhibition measures). The final data set consisted of 249 participants. All participants completed the focal and nonfocal cue blocks (order counterbalanced), followed by symmetry span, AX-CPT verbal, task-switching verbal, sustained attention visuospatial, AX-CPT visuospatial, reading span, sustained attention verbal, and task switching visuospatial. Participants completed all tasks in group sessions in a single laboratory session that lasted approximately two hours.

Materials and procedure.

PM task. This version of the task was similar to Experiment 3 except only had delays of 0, 1,000, and 1,600 ms. There were six cues presented at each delay, with each delay being presented twice in each third of the block. To account for the increased number of cues, each block consisted of 465 trials. In the focal block the PM cue was the word *sister*. In the nonfocal block, the PM cue was any word starting with "S" (spleen, smoke, sugar, sand, silver, stereo, shoulder, spring, skate, speaker, saved, self, senator, senior, slot, snail, sport, surface).

Reading span. Participants were required to read sentences while trying to remember a set of unrelated letters. For this task, participants read a sentence and determined whether the sentence made sense or not (e.g., "The prosecutor's dish was lost because it was not based on fact?"). Half of the sentences made sense while the other half did not. Nonsense sentences were made by simply changing one word (e.g., "dish" from "case") from an otherwise normal sentence. Participants were required to read the sentence and to indicate whether it made sense or not. After participants gave their response, they were presented with a letter for 1,000 ms. At recall, the letters from the current set were recalled in the correct order by clicking on the appropriate letters. There were two trials of each list-length with the list-length ranging from three to seven letters.

Symmetry span. In this task, participants were required to recall sequences of red squares within a matrix while performing a symmetry-judgment task. In the symmetry-judgment task participants were shown an 8×8 matrix with some squares filled in black. Participants decided whether the design was symmetrical across its vertical axis. The pattern was symmetrical half of the time. Immediately after determining whether the pattern was symmetrical, participants were presented with a 4×4 matrix with one of the cells filled in red for 650 ms. At recall, participants recalled the sequence of red-square locations in the preceding displays, in the order they appeared, by clicking on the cells of an empty matrix. There were two trials of each list-length with the list-length ranging from two to five squares. The dependent variable for both tasks was the proportion of correct items in the correct position.

Task switching. In this task the participant must correctly identify whether a target stimuli fits into one of two categories and contains three unique blocks. In the verbal version of the task, during the first block the participant must identify whether a value is greater (values: 6, 7, 8, and 9) or less (values: 1, 2, 3, and 4) than 5. The participant presses the "Q" key if the value is less than 5 and the "P" key if it is greater than 5. The participant completes eight practice trials during which they are given feedback to their performance. After the practice trials, they complete 40 experimental trials. After completing the previous block, they are then given new identification criteria. Specifically, they identify whether the value shown is odd ("Q" key) or even ("P" key). The progression of the block and number of trials are identical to the previous block (i.e., practice with feedback then experimental trials). Next, they complete a block where the identification criteria switches every three trials (i.e., two trials of value, two trials of odd/even, two trials of value, etc.). Participants complete 16 (16) practice trials and then 80 (80) experimental trials. As before, the participant presses the "Q" key if the value is less than 5 and the "P" key if it is greater than 5, and the "Q" key for odd values and "P" key for even values.

The visuospatial version of task switching has the same overall structure but with different stimuli and identification criterion. As before, each block begins with practice trials followed by experimental trials. During the first block they are shown letters (uppercase: "A", "B", "F", and "H"; lowercase: "d", "e", "r", and "t") and tasked with identifying the case of the letter ("Q" key for uppercase and "P" key for lowercase). During the second block they identify what color the letter is written in. The stimuli are the same letters used from the first block, and the two ink colors are red and blue. To identify a letter written in red they press the "Q" key and for a letter written in blue they press the "P" key. During the third block the identification criterion switches every third trial. The dependent variable for both tasks was accuracy switch trials.

Sustained attention. For the verbal version of this task participants are shown a digit (Numbers 1–9) in the center of the screen. The participant must respond to every digit except the digit 3. The digit appears on the screen briefly (200 ms) followed by a mask, which remains on the screen for 900 ms. The participant makes their response ("spacebar") while the mask is on the screen. If they fail to respond within 900ms the trial is coded as incorrect. The participant completes 18 practice trials with feedback then completes 144 experimental trials where no performance feedback is given. Each stimulus is presented an equal number of times.

For the visuospatial version of this task the overarching structure is the same, but the stimulus is changed. Specifically, they are shown colored (pink, brown, blue, gray, green, orange, purple, red, and yellow) squares. They respond ("spacebar") to all colors except blue. The dependent variable for both tasks was the number of false alarms (i.e., pressing "spacebar") to "no-go" trials (i.e., the number 3, or blue squares).

AX-CPT. An adapted version of this task was constructed based upon the Redick and Engle (2011) version. A letter (any letter except X, K, or Y) was presented in the center of the screen for 1,000 ms at the beginning of each experimental trial. After the stimulus appears the participant was instructed to use the number pad on the keyboard to press the "1" key with their pointer finger. An unfilled interstimulus interval of 2,000 ms followed. After the interstimulus interval, the probe appeared. The probe was a letter (any letter except A, K, or Y) presented in the center of the screen for 500 ms. Following the probe, a triangle shape made by three "+" symbols appeared during the 1,000 ms intertrial interval. Participants were instructed to press the target button ("2" key) with the middle finger of their right hand as quickly as possible whenever they observed an "A" cue followed by an "X" probe, and to press the nontarget key with the index finger of the right hand ("1" key) as quickly as possible whenever they observed any other letter pair. Participants were instructed to respond only once they observed the second letter in the pair (i.e., the probe). Responses to the probe stimuli were recorded with a time limit of 1,500 ms. When participants responded incorrectly a tone was played through their headphones. On correct trials there was no tone played.

The proportions of trial types were based on those used by Richmond, Redick, and Braver (2015): 40% of the trials in each task block consisted of an "A" followed by an "X" (AX trials), 10% of the trials in each block consisted of an "A" followed by a letter other than "X" (pseudorandomly selected; AY trials), 10% of the trials in each block consisted of a letter other than A (pseudorandomly selected) followed by an "X" (BX trials), and 40% of the trials in each block consisted of a letter other than "A" (pseudorandomly selected) followed by a letter other than "X" (pseudorandomly selected; BY trials). Trials within each block were presented randomly. Participants complete 6 practice trials of which they must get 70% correct before proceeding to experimental trials (i.e., practice trials repeats until 70% correct is achieved). In addition, there are two blocks of 80 experimental trials (160 experimental trials total).

For the visuospatial version of AX-CPT the overall structure of the task is the same as the verbal version but the stimuli are dot patterns (Ball & Brewer, 2018). The participants are given a stimuli pairing (like AX) that they need to respond to. The dependent variable for both tasks was the number of AX hits minus BX false alarms (Ball & Brewer, 2018).

Procedure. After participants consented to participate in the experiment, participants completed the tasks in the following order: PM, symmetry span. AXCPT verbal, task switching verbal, sustained attention visuospatial, AXCPT visuospatial, reading span, sustained attention verbal, tasking switching verbal.

Results

Overall performance (without cognitive ability). Late PM responses (within two trials, $M = .032$, $SE = .002$) were rare and counted as correct. Cue detection, ongoing task accuracy, and ongoing task RTs were submitted separately to a 2 (cue type: focal vs. nonfocal; within-subjects) × 3 (delay: 0 vs. 1,000 vs. 1,600 ms; within-subjects) ANOVA. Results can be found in Figure 5.

Cue detection. Cue detection was higher for focal cues [cue type: $F(1, 248) = 283.65$, $p < .001$, $\eta_p^2 = .534$] and increased with delays [delay: $F(2, 496) = 30.81$, $p < .001$, $\eta_p^2 = .110$]. Critically, the focality effect *was* reduced with delays [Cue Type × Delay: $F(2, 496) = 13.43$, $p < .001$, $\eta_p^2 = .051$]. Planned contrasts revealed that focal cue detection was higher than nonfocal cue detection at all delays (all ps $< .001$).

Ongoing task performance.

RT. RT was faster in the focal block [cue type: $F(1, 248) = 5.66$, $p = .018$, $\eta_p^2 = .022$] and decreased with delay [delay: $F(2, 496) = 3838.57$, $p < .001$, $\eta_p^2 = .939$]. Critically, the faster RT in the focal block changed across delays [cue type x delay: $F(2, 496) = 6.40$, $p = .002$, $\eta_p^2 = .025$]. The Cue × Delay interaction reflects that focal RTs were faster than nonfocal RTs with no delay ($p = .001$), but not with delays of 1,000 and 1,600 ms (p*'s* $> .229$).

Accuracy. Accuracy increased with delay [delay: $F(2, 496) = 56.07$, $p < .001$, $\eta_p^2 = .184$]. However, there were no other significant effects. Accuracy was not better in the focal block [cue type: $F < 1$] and this did not change across delays [Cue Type × Delay: $F < 1$].

Overall performance (with cognitive ability). Tasks within a construct were generally more highly correlated than tasks between constructs, indicating reasonable convergent and discriminant validity. A principal components analysis was separately conducted on the working memory, proactive control, inhibition, and task switching measures to create constructs to assess individual differences in performance. Cue detection was submitted to a

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

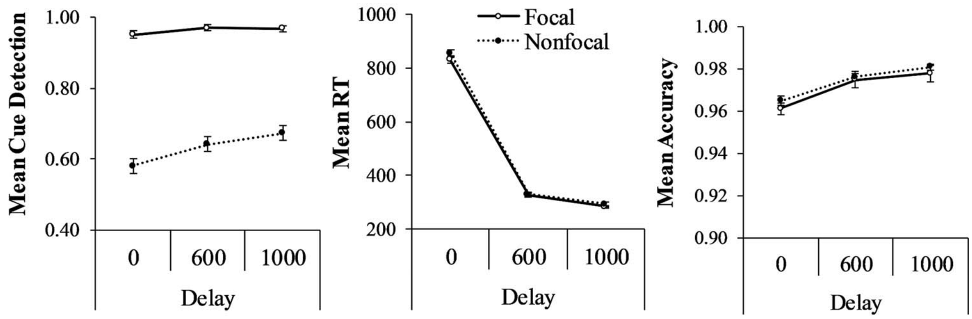

Figure 5. Mean cue detection (left panel), ongoing task reaction times (RTs; middle panel), and ongoing task accuracy (right panel) as a function of cue type (focal vs. nonfocal) and delay in Experiment 5. Error bars reflect standard errors.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

2 (cue type: focal vs. nonfocal; within-subjects) × 3 (delay: 0 vs. 1,000 vs. 1,600 ms; within-subjects) generalized linear model with cognitive ability (e.g., WMC) entered as a covariate. Note that analyses were conducted on the entire sample of participants, whereas Table 1 displays performance for individuals in the upper and lower 25th percentiles (except for the inhibition task which was split at the 50th percentile). Figures are for illustrative purposes only.

Working memory capacity. Overall, somewhat surprisingly, those with higher working memory ability did not detect more cues [WMC: $F(1, 247) = 3.64$, $p = .058$, $\eta_p^2 = .015$]. Working memory did not interact with cue type [WMC × Cue Type: $F(1, 247) = 1.05$, $p = .307$, $\eta_p^2 = .004$], or delay [WMC × Delay: $F(2, 494) = 2.63$, $p = .073$, $\eta_p^2 = .011$]. Although WMC did moderate the focality effect with increased delay [WMC × Cue Type × Delay: $F(2, 494) = 5.97$, $p = .003$, $\eta_p^2 = .024$], it was not expected from the delay theory. The moderating effect of WMC was primarily driven by the *improvement* for high WMC participants from 0 to 1,000 ms. That is, the delay theory would predict that the WMC difference would be largest at 0 ms and decrease across delays (i.e., at 1,000 ms).

Proactive control. Overall, those higher in proactive control ability detected more cues [proactive control: $F(1, 247) = 33.19$, $p < .001$, $\eta_p^2 = .118$]. This proactive control advantage was more pronounced in the nonfocal block [Proactive Control × Cue Type: $F(1, 247) = 16.52$, $p < .001$, $\eta_p^2 = .063$], but not with delay [Proactive Control × Delay: $F < 1$]. Critically, PC did not moderate the focality effect with increased delay [Proactive Control × Cue Type × Delay: $F < 1$].

Inhibition. Overall, those higher in inhibition ability detected more cues [inhibition: $F(1, 247) = 10.24$, $p = .002$, $\eta_p^2 = .049$]. This inhibition advantage did not differ by cue type [Inhibition × Cue Type: $F(1, 247) = 3.37$, $p < .068$, $\eta_p^2 = .013$], or delay [Inhibition × Delay: $F < 1$]. Critically, inhibition did not moderate the focality effect with increased delay [Inhibition × Cue Type × Delay: $F(1, 247) = 1.03$, $p = .357$, $\eta_p^2 = .004$].

It should be noted, however, that accuracy on the inhibition measures was extremely high. In fact, 177 of the 249 participants committed no errors in either task. Because of the high performance, the "high" and "low" inhibition ability participants displayed in Table 1 is actually based on a split between those that did or did not commit errors (rather than a quartile split like the other

Table 1
Prospective Memory Performance in Experiment 5 for Individuals Low and High in Cognitive Abilities

Variable	Focal 0 ms	Focal 1,000 ms	Focal 1,600 ms	Nonfocal 0 ms	Nonfocal 1,000 ms	Nonfocal 1,600 ms
Working memory						
Low	0.92 (0.02)	0.95 (0.02)	0.95 (0.02)	0.59 (0.04)	0.58 (0.04)	0.69 (0.04)
High	0.97 (0.02)	0.97 (0.02)	0.98 (0.02)	0.59 (0.04)	0.70 (0.04)	0.69 (0.04)
Proactive control						
Low	0.92 (0.02)	0.94 (0.02)	0.95 (0.02)	0.46 (0.04)	0.52 (0.04)	0.55 (0.05)
High	0.98 (0.02)	0.98 (0.02)	0.98 (0.02)	0.70 (0.04)	0.76 (0.04)	0.78 (0.04)
Inhibition						
Low	0.96 (0.01)	0.97 (0.01)	0.96 (0.02)	0.49 (0.04)	0.54 (0.04)	0.60 (0.04)
High	0.95 (0.01)	0.97 (0.01)	0.97 (0.01)	0.60 (0.02)	0.67 (0.02)	0.70 (0.02)
Task switching						
Low	0.87 (0.03)	0.93 (0.03)	0.90 (0.03)	0.51 (0.04)	0.52 (0.05)	0.60 (0.04)
High	0.97 (0.01)	0.98 (0.01)	0.98 (0.01)	0.65 (0.03)	0.69 (0.03)	0.69 (0.03)

Note. Low and high ability, respectively, refer to participants who scored in the lower and upper 25th percentile for all cognitive abilities except inhibition. For inhibition, low and high ability reflect the lower and upper 50th percentile. Values in parentheses are standard errors.

ability measures). This measure may not necessarily be a sensitive indicator of inhibitory processes and the results should be interpreted with caution.

Task switching. Overall, those higher in task switching ability detected more cues [task switching: $F(1, 247) = 11.74$, $p = .001$, $\eta_p^2 = .045$]. This task switching advantage did not differ by cue type [Task Switching × Cue Type: $F(1, 247) = 1.45$, $p = .230$, $\eta_p^2 = .006$], or delay [Task Switching × Delay: $F < 1$]. Critically, task switching did not moderate the focality effect with increased delay [Task Switching × Cue Type × Delay: $F(2, 494) = 2.96$, $p = .053$, $\eta_p^2 = .012$].

Discussion

Although a bit more nuanced than the previous experiments, the results of Experiment 5 are fairly clear. For the first time, the anticipated reduction in the focality effect with increased delays was evident. Moreover, delays reduced the ongoing task RT differences typically seen between focal and nonfocal conditions (consistent with Experiments 1 and 2). The significant effects on both cue detection and ongoing task responding may reflect the increased power to detect small effects given the large sample size and increased number of cue presentations. These results most closely parallel the original reported findings from Loft and Remington (2013). Notably, as with all previous experiments the focality effect was never completely eliminated even at the longest delay (see also Loft et al., 2014).

More interestingly, there was converging evidence across multiple measures that attention ability did *not* moderate the reduction in the focality effect with increased delays. These findings seem difficult to reconcile from a delay mechanism account of cue detection. One possible reason that low ability participants do more poorly at nonfocal cue detection is because they do not set high enough response thresholds during the 0-ms delay condition, so ongoing task information accumulates prior to PM information (Strickland et al., 2017). The 1,000 and 1,600 ms tone exogenously causes one to increase their response threshold by forcing participants to withhold responses. This procedure therefore circumvents the typical suboptimal threshold settings that low ability participants employ (i.e., during 0-ms delay), which in turn should allow more time for PM information to accumulate and substantial increases to cue detection. Delay theory therefore predicts that low ability participants should benefit more from delays, as high ability participants already set optimal thresholds and should not benefit much from delayed responding. Inconsistent with this idea, cognitive ability did not moderate the effect of delay, although all participants did show some benefit to nonfocal performance with delays. These results are, however, consistent with predictions from monitoring theory. Monitoring theory posits that low ability participants are less able to sustain monitoring across trials due to capacity limitations and consequently are more likely to miss PM cues. Across three of the four cognitive abilities assessed, high ability participants outperformed low ability participants, regardless of delay. Together these findings suggest that cognitive ability, but not response thresholds, are most critical for successful cue detection.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Meta-Analysis

The results of the current study are somewhat at odds with the previous studies showing an effect of delay on the focality effect (Loft & Remington, 2013; Loft et al., 2014). Given the apparent discrepancies, we conducted a meta-analysis that included the previous and current effect sizes. This analysis was conducted to determine whether the magnitude of the focality effect (i.e., focal–nonfocal PM) was reduced at the longest delay (e.g., 1,600 ms) relative to the shortest delay (e.g., 0 ms). Separate analyses for the focality difference at the short and long delays can be found in the online supplemental material. Table 2 describes the characteristics of each study entered into the meta-analysis.[3] Analyses were conducted with meta-essentials software (Hak, Van Rhee, & Suurmond, 2016). A random effects model was applied to account for the heterogeneity across studies, using inverse variance weighting of the effect sizes (i.e., larger confidence intervals had less weight).

The forest plot in the top half of Figure 6 shows the mean weighted effect size (circles), 95% confidence interval (CI; solid line), and 95% prediction interval (PI; dashed line). CIs reflect the range with which the true effect lies, whereas PIs reflect the range of predicted true treatment effects for future studies. As can be seen for the combined analysis, the CI does not overlap with zero indicating that the focality effect is indeed reduced with increased delays. This is consistent with predictions of delay theory. However, the PI does overlap with zero indicating that future studies should anticipate some null effects. Inconsistent with predictions of the delay theory, however, was the supplemental finding that the focality effect was not completely eliminated at the longest delay (see the online supplemental materials).

It is important to note that there appears to be two different "true" population effect sizes (blue circles): one from the current set of studies (i.e., Ball subgroup) and another from the previously published studies (i.e., Loft subgroup). Given the heterogeneity in the combined effect ($I^2 = 59\%$), a subgroup analysis was performed with weighting averaged separately for each subgroup (Borenstein, Hedges, Higgins, & Rothstein, 2009). Consistent with the noted discrepancies between the current and previous studies, this analysis[4] indicated that the combined effect size is significantly greater for the Loft subgroup than the Ball subgroup, $Q = 5.55$, $p = .018$. As can be seen visually in the subgroup plots, both the CI and PI overlap with zero in the Ball subgroup, but neither do in the Loft subgroup.

The funnel plot in the bottom half of Figure 6 reports the distribution of effect sizes across experiments. Funnel plots are often used to assess for study heterogeneity, reporting bias, and chance by examining the distribution of effect sizes around the meta-analytic estimate of the effect size (Sterne et al., 2011). In meta-analyses of studies with minimal heterogeneity and no re-

[3] It is important to note that many of the experiments used different delay intervals. For example, in the current study Experiment 1 included delays of 0, 600, 1,000, 1,600, whereas Experiment 5 only used 0, 1,000, 1,600. Likewise, Loft and Remington (2013) used delays of 400, 600, 1,000, and 1,600 (Experiment 1), or 0, 200, 400, and 600 (Experiment 2). The meta-analysis therefore only includes the shortest (e.g., 0 ms) and longest delay (e.g., 1,600 ms) intervals, up to 1,600 ms, for each study.

[4] There were no differences between subgroups defined by other study characteristics listed in Table 2 (e.g., nonfocal cue type).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 2
Study Characteristics and Descriptives for Each Experiment in the Meta-Analysis

	Study characteristics						Focality effect							
Manuscript and experiment	Short interval	Long interval	Nonfocal cue type	Nonfocal cue exposure	Tone onset	Sample	*N*	Short	Long	*g*	*SE*	CI	PI	I^2
Ball et al. (current): [B]														
1	0	1,600	Syllable	Unique	Variable	Healthy	40	0.15	0.23	−0.16	0.11	[−0.16, −0.39]	—	—
1	0	1,600	Syllable	Repeated	Variable	Healthy	37	0.18	0.15	0.07	0.14	[−0.21, 0.35]	—	—
2	0	1,600	Syllable	Unique	From 0	Healthy	43	0.15	0.13	0.06	0.14	[−0.23, 0.35]	—	—
2	0	1,500	Syllable	Repeated	From 0	Healthy	40	0.15	0.19	−0.09	0.18	[−0.46, 0.29]	—	—
3	0	1,600	First letter	Unique	From 0	Healthy	40	0.31	0.16	0.38	0.19	[−0.01, 0.76]	—	—
4	0	1,500	First letter	Unique	From 0	Healthy	41	0.39	0.22	0.43	0.19	[0.04, 0.81]	—	—
5	0	1,600	First letter	Unique	From 0	Healthy	249	0.37	0.30	0.22	0.04	[0.13, 0.31]	—	—
Loft & Remington (2013): [LR]														
1	400	1,600	Category	Unique	Variable	Healthy	72	0.16	0.03	0.32	0.09	[0.14, 0.49]	—	—
2	0	600	Category	Unique	Variable	Healthy	72	0.17	0.06	0.29	0.12	[0.05, 0.52]	—	—
3	0	1,600	Syllable	Unique	Variable	Healthy	72	0.27	0.07	0.40	0.11	[0.18, 0.63]	—	—
Loft, Bowden, Ball, and Brewer (2014): [L]														
1	0	1,600	Syllable	Unique	Variable	HIV+	57	0.32	0.15	0.42	0.15	[0.11, 0.72]	—	—
Combined	—	—	—	—	—	—	**764**	**0.24**	**0.15**	**0.21**	**0.06**	**[0.08, 0.34]**	**[-0.11, 0.53]**	**59%**
Ball subgroup	—	—	—	—	—	—	*490*	*0.24*	*0.20*	*0.12*	—	*[−0.04, 0.27]*	*[−0.31, 0.54]*	*62%*
Loft subgroup	—	—	—	—	—	—	*273*	*0.23*	*0.08*	*0.34*	—	*[0.29, 0.40]*	*[0.25, 0.44]*	*0%*

Note. *N* = sample size; *g* = Hedge's *g* effect size; *SE* = standard error of effect size; CI = confidence interval of effect size; PI = prediction interval of effect size; I^2 = heterogeneity of effect size. Numbers in bold reflect the combined effects across all studies; numbers in italic reflect the effects separately for each subgroup.

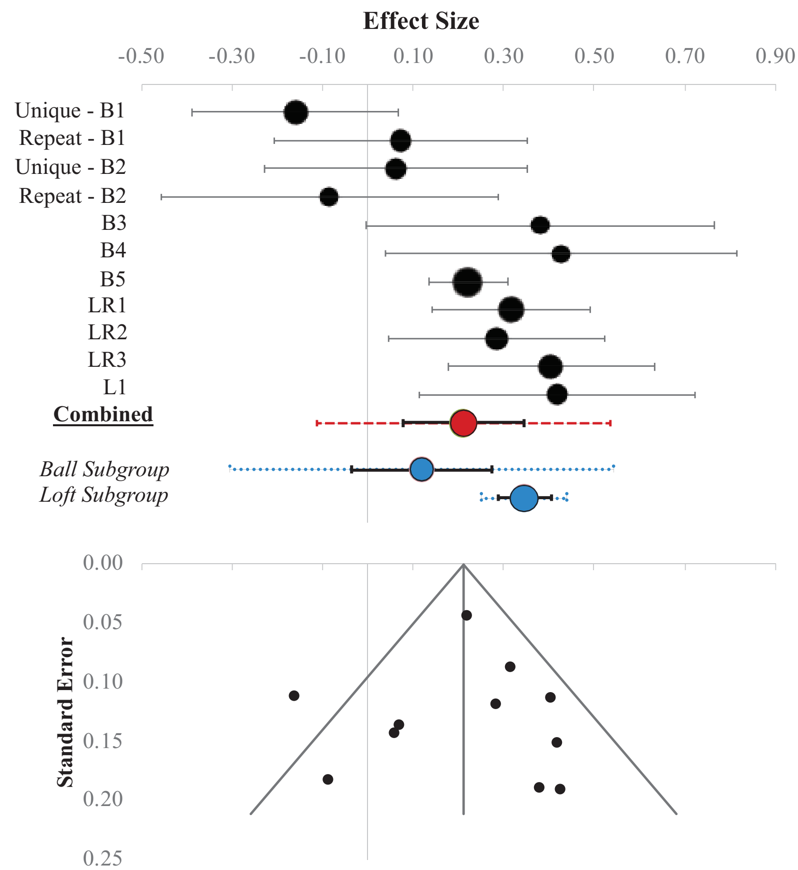

Figure 6. The forest plot (top) shows the mean weighted effect size (circles) 95% confidence interval (CI; solid line), and 95% prediction interval (PI; dashed line). Effects are reported for each individual experiment (black circle), all combined (long dashed line), and separately for the current and previous studies (short dashed lines). The funnel plot (bottom) reports the distribution of effect sizes across experiments. B = Ball et al. (current article); L = Loft et al. (2014); LR = Loft and Remington (2013). The numbers refer to the experiment in each article. "Unique" and "Repeat" refer to the between-subjects manipulation of Experiment 1 and 2 of the current article. See the online article for the color version of this figure.

porting bias one should expect that effect sizes randomly vary around the pooled effect size (horizontal axis) with studies having more power to have smaller standard errors (vertical axis). Thus, the distribution of effect sizes should resemble an inverted funnel and the triangle centered on the meta-analytic effect size represented in Figure 6 shows reflects 1.96 standard errors on both sides so the reader can examine whether the study effect sizes are randomly distributed and whether any study effect size falls outside of this interval. The funnel plots in Figure 6 show a reasonably random distribution of effect sizes around the meta-analytic effect size (effect size = .21) with only one study (Ball 1; Unique) falling outside of the range. The Egger regression was not significant, $t = 0.08$, $p = .941$, indicating no publication bias.

General Discussion

Previous research suggests that delays as short as 600 ms eliminate the typical finding of higher focal cue than nonfocal cue detection (Loft & Remington, 2013). Those results provided the foundation for the delay theory of PM (and later the PMDC model), a new and exciting model of PM that challenges traditional views of PM processing. Given the importance of such findings, the current study examined the generalizability of the delay mechanism. Unfortunately, however, the results were far from conclusive in supporting those original findings. Only one experiment (with over 200 people) showed the anticipated reduction in the focality effect across the entire range of delays, and the effect was never eliminated even when participants had over 2 s to respond. Moreover, testing predictions from competing theories of PM failed to provide support for the core theoretical assumptions of delay theory. Such findings call into question the explanatory power of a threshold mechanism in supporting cue detection. Below we discuss theoretical implications of competing theories of PM.

Generalizability

The primary goal of the current study was to examine the viability of a delay mechanism in explaining performance. To test this idea, we used a tone procedure similar to Loft and Remington

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

(2013) to examine whether delay reduces or eliminates the focality effect. It is important to note, however, that none of our experiments consisted of direct replication of the exact procedure and sample size used in the original study. Experiment 1 used the same tone procedure and cues as the original procedure but did not include a control (no intention) block and the "unique cues" condition had only 40 participants compared to 72 of Loft and Remington. However, it is not entirely clear why having a control block randomly intermixed within the procedure would influence performance and the total sample size (N = 77) collapsed across the stimulus exposure factor (unique vs. repeated) was comparable to the original study. These issues aside, somewhat surprisingly, this experiment failed to show that delays reduced the focality effect despite improving ongoing task responding in the nonfocal block. Experiment 2 used the same PM cues as Experiment 1 but modified the tone procedure that signaled delays. Despite this change, Experiment 2 showed a nearly identical pattern of results to Experiment 1 and failed to find a reduction in the focality effect. These experiments call into question the generalizability of a delay mechanism in contributing to PM performance.

Given that we were unable to reproduce the original findings using the same cues and delay intervals, Experiments 3–5 assessed performance with different nonfocal cues (first letter) and combinations of delays (up to 2,500 ms). Although Experiments 3 and 4 showed no statistical reduction in the focality effect with increased delays, the results numerically trended in that direction. It is possible with more power we would have found a significant interaction effect. The high-powered individual differences study in Experiment 5 did find the anticipated reduction in the focality effect. These findings provide support for the original findings demonstrating that the focality effect may be reduced by forced delays and echo the sentiments originally raised by Loft and Remington (2013) that the type of nonfocal cue (e.g., syllable vs. first letter) appears to influence the degree with which a delay may benefit nonfocal PM performance. However, none of the experiments showed a complete elimination of the effect, even with delays as long as 2,500 ms.

Together with Loft and Remington (2013) and Loft et al. (2014), this means that five out of nine different experiments have found the reduction in the focality effect with increased delays. Given the discrepancies across the current and previous studies, we performed a meta-analysis to examine the combined effect across all studies. Consistent with the delay theory, there was a significant, albeit small, reduction in the focality effect in the longest (e.g., 1,600 ms) relative to the shortest (e.g., 0 ms) delay. This effect, however, appeared to be largely driven by effect sizes reported in prior research rather than effect sizes from the current research. Given the small combined effect size (g = .22) and that the lower bound of the CI and PI respectively approached (g = .08) and encompassed zero (g = −.11), this warrants cautious optimism in regard to demonstrating a reduction in the delayed focality effect in future studies. The one finding that does provide clear evidence against delay theory predictions is that the focality effect was not completely eliminated at the longest delay (see the online supplemental material). This represents an important boundary condition for future theorizing.

Finally, it is worth noting that ongoing task performance largely replicated the findings from Loft and Remington (2013). Across all experiments, ongoing task accuracy and response times improved with delay. This indicates that participants were processing lexical information (and presumably cue-relevant features) during the delay and preparing for a response. Moreover, in Experiments 1, 2, and 5 the difference between focal and nonfocal RTs was reduced with delay. This suggests that the methodology used in the current study was at least conceptually similar to the original studies. Why participants were unable to use the delays to similarly improve PM, however, is not entirely clear.

Challenges to the Delay Mechanism

The fact that we were unable to directly replicate all the findings of Loft and Remington (2013) across all experiments is not necessarily a concern, nor was it our intention to do so. As described above our methodology differed slightly (Experiment 1 and 2) to more considerably (Experiment 3–5) from the original procedure. The purpose of the study was to better understand how or if a delay mechanism could explain focality effects typically observed in PM. What is concerning from a theoretical perspective is the finding that delay had *no* influence on PM performance in Experiments 1 and 2 and only a minimal influence in Experiments 3 and 4. The delay theory suggests that delays, whether implemented exogenously by a tone or endogenously by a participant, should allow more time for the PM response to race and compete for response selection with the more routine ongoing task response. This was clearly not the case in Experiments 1 and 2, and it is not clear why the delay mechanism would operate for some cues (e.g., using "first letter S" in Experiment 5) but not others (e.g., using "TOR syllable" in Experiment 1). Another concern from a theoretical standpoint is that the effect was never eliminated, even at relatively long delays (e.g., 2,500 ms in Experiment 4). If delays are the sole contributor to performance, then given enough time nonfocal cue detection should always approximate focal cue detection and both should reach ceiling. In reality, however, as can be seen in Experiment 4 there appears to be an upper limit (or perhaps a point of diminishing returns), whereby after 1,500–1,600 ms, participants no longer show any benefit to performance. Intuitively this makes sense, as at some point mind wandering or boredom may set in. But without the addition of such mechanism or a stopping point added to the existing delay mechanism, the delay theory simply cannot account for such findings.

Also particularly difficult to reconcile from a delay theory perspective is the finding that low attention ability in younger adults did not show a reduction in the focality effect with increased delays. Considerable research with younger adults has demonstrated that those higher in attention control ability detect more nonfocal cues than low ability participants (Ball, Vogel, & Brewer, 2019; Ball & Brewer, 2018; Brewer et al., 2010). To explain these findings, proponents of the delay theory have argued that those lower in attention control ability may adopt a more liberal response threshold in the PM block that does not allow sufficient time for PM response selection to occur (Strickland et al., 2017). The logic then follows that by forcing low ability younger adults to withhold responding this should circumvent any difficulties these individuals typically display in setting appropriate response threshold, thereby providing considerable improvements to performance. In contrast, because high ability participants are already setting appropriate response thresholds, forcing them to withhold responding should not produce as large of a benefit to performance. The

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

results of Experiment 5 clearly do not support this idea, as high and low ability younger adults showed similar (minimal) improvements from delay. Critically, however, high ability participants detected more cues than low ability participants, which is consistent with the majority of extant research. Again, these findings can be accommodated by monitoring theory: low ability participants have greater difficulty in actively maintaining the intention that results in worse overall cue detection, and because withholding ongoing task responding is independent of active maintenance of the intention, these individuals do not show any additional benefits to performance with increased delays.

It is important to note that the processes measured using the tone procedure do not necessarily directly map onto those measured by evidence accumulation models from which the delay theory was formally derived. However, Anderson et al. (2018) formally tested delay theory assumptions using such modeling techniques. In that experiment participants were instructed to "delay" their responding during the PM block by adopting a more stringent criteria for making their ongoing task decisions. As anticipated, these instructions resulted in increased response thresholds relative to standard PM instructions. Despite increased thresholds in the delay condition, cue detection was equivalent between the delay and standard PM conditions. Thus, it appears that an actual delay strategy does not map onto the purported parameter thought to reflect delayed responding. Taken together with the results of the current study, these findings suggest that the delay theory is an incomplete theory of PM processing.

Alternative Mechanisms

Monitoring theory. Although these results do not appear compatible with the delay theory, such findings can be readily accommodated by monitoring theories of PM. The primary mechanism underlying monitoring theories of PM is active maintenance of the PM intention while busily engaged in ongoing task processing. Given that some available capacity is devoted to noticing cues, this reduces processing resources that would otherwise be devoted to ongoing task processing. Within the different theories, PM failures can occur because bottom-up processes are not sufficient for intention retrieval or because top-down monitoring is not actively engaged (dynamic multiprocess framework; Scullin et al., 2013; Shelton & Scullin, 2017), a prospective retrieval mode or target checking is not initiated (two-stage checking hypothesis; Guynn, 2003; Guynn, McDaniel, & Einstein, 2001), preparatory attention is not sustained or recognition checks fail (preparatory attentional and memory theory; Smith, 2003), metacognitive assessments of task difficulty are miscalibrated (attention allocation hypothesis; Hicks, Marsh, & Cook, 2005), or intention maintenance waxes and wanes across time (periodic reminding view; Ball, Brewer, Loft, & Bowden, 2015; Dewitt, Knight, Hicks, & Ball, 2012; Einstein, McDaniel, Williford, Pagan, & Dismukes, 2003). Although we did not find strong evidence for delays influencing performance, mechanistically it makes sense that unfilled delay intervals could be used to engage in some sort of monitoring process (e.g., target checking). Thus, delays may facilitate performance not by allowing more time for PM to accumulate but rather by reducing ongoing task demands that increases the likelihood that a check is made to determine if a PM cue is present. If low attention ability participants have greater difficulty in actively maintaining this intention, then reduced demands should not show any additional benefit to performance with decreased delays. Assuming that attentionally demanding monitoring (e.g., target checking) processes may periodically lapse across (or even within) trials, then it is reasonable to assume that delays could also decrease nonfocal cue detection. For example, with delays longer than 2 s, participants may begin to mind wander, nullifying the potential benefit of an unfilled delay for engaging in a target check. This latter idea is similar to delay-execute PM research showing that rapid forgetting occurs when the intention is retrieved but response execution must be temporarily delayed (e.g., 2 s; Ball, Knight, Dewitt, & Brewer, 2013; Einstein et al., 2003). Importantly, however, the delay theory would anticipate a monotonic increase with delays rather than a point of diminishing returns.

Mechanistically, Ball and Brewer (2018) have argued that proactive and reactive control, respectively, may underlie nonfocal and focal cue detection (see also Ball, 2015; Braver, 2012; Bugg, McDaniel, & Einstein, 2013; Bugg, McDaniel, Scullin, & Braver, 2011). Proactive control involves top-down engagement of attention in a preparatory fashion (i.e., preparatory monitoring), whereas reactive control involves the bottom-up engagement of attention in a transient manner following stimulus presentation (i.e., spontaneous retrieval). Across several studies it was found that those higher in proactive control ability detected more nonfocal cues. Moreover, those individuals exhibited greater slowing in the mu parameter derived from ex-Gaussian analyses, which is thought to reflect slowing on each trial associated with checking for PM cues (Ball et al., 2015; Loft et al., 2014). It was suggested that those high in proactive control ability were more likely to sustain monitoring and inhibit prepotent ongoing task responding on each trial to check for PM cues. The results of the current study are consistent with this idea, showing that high proactive control individuals detected more cues than low ability participants across all delays.

Decision control theory. The results can also be accommodated by the more recent PMDC model (Boag, Strickland, Heathcote, et al., 2019; Boag, Strickland, Loft, & Heathcote 2019; Strickland et al., 2018; Strickland et al., 2019). Although delay theory is a single process model (threshold only), PMDC is a multiprocess model of PM and incorporates the concepts of proactive and reactive control processes to account for PM processing. In addition to the proactive setting of response thresholds prior to each trial, the PMDC model posits that participants reactively adjust accumulators following stimulus onset. This reactive control can come in the form of excitation of PM accumulators (e.g., speeding accumulation of PM information) or inhibition of ongoing task accumulators (e.g., slowing accumulation of non-PM information). Interestingly, reactive inhibition in the PMDC model is conceptually similar to our (Ball & Brewer, 2018) interpretation of proactive control in that both suggest the underlying mechanisms contributing to cue detection is the inhibition of ongoing task processing.

Strickland et al. (2018) fit this model to PM data and found that both ongoing task and PM thresholds were greater in nonfocal than focal blocks, but surprisingly these threshold changes were not predictive of PM performance. Although this finding seems contradictory to the core assumptions of the delay theory of PM that suggest that response thresholds are critical for prospective remembering, it is consistent with the results of the current study showing that delay had little influence on performance. In contrast, on focal and nonfocal PM cue trials there was greater reactive inhibition of ongoing task (word and nonword) accumulators, which was predictive of performance, along with greater excitation of the PM accumulator. Based on these findings, it was suggested that cue-specific inhibitory processes may be the primary driver of PM performance rather than a delay mechanism. The finding that reactive inhibition predicted performance is generally consistent with the current study and those of Ball and Brewer (2018) showing that high proactive control ability participants detected more cues than low ability participants. The PMDC model could account for the current findings by suggesting that forced delays did not influence reactive inhibition and, therefore, produced little to no influence on PM. Based on this interpretation, the results from Experiment 5 and previous research showing a relationship between attention ability and PM (e.g., Ball et al., 2018; Brewer et al., 2010; Schnitzspahn et al., 2013; Zuber et al., 2016) might reflect that control processes are needed specifically on cue trials (Bugg et al., 2013).

Remaining issues. Although monitoring (qualitative) and PMDC (quantitative) theories of PM can generally account for the current data, they are not without their limitations. In monitoring theory, it is easy to invoke ideas of "waxing/waning" of attention or "target checks" failing. How to empirically quantify these processes is the primary issue that such verbal theories face, and delay theory rightly calls into question the utility of such explanatory variables. The PMDC model provides some resolution to this issue by using a mathematically tractable model that can constrain theory based on a priori predictions. The issue is that the cognitive processes (e.g., processing efficiency, response caution, motor responses) associated with the parameters (e.g., drift rate, thresholds, nondecision time) derived from these models having nothing to do with PM processing. In fact, the cognitive interpretation of PM-related "proactive" or "reactive" control can only be applied outside of the decision process (i.e., during nondecision time). Applying these labels to model parameters in the context of PM results in the same issue as existing verbal theories of PM. Although we see no immediate resolution to these issues and debates, we believe that continued research combining experimental, mathematical, physiological, and individual differences methodologies can lead to the development of more tractable process-based models of PM.

Conclusions

One's ability to detect PM cues and retrieve appropriately planned behaviors in the future must be dependent on a diverse array of cognitive processes that incorporate both attention and long-term memory mechanisms (Ball et al., 2019). In the current study we sought to provide additional empirical and theoretical support for novel decision-making mechanisms proposed by delay theory (Heathcote et al., 2015). As described previously, we believe that the Loft and Remington (2013) results are some of the most interesting and theoretically important PM findings in recent literature. Altogether, however, the results from the current study failed to provide adequate support for this theory. Instead our findings highlight that, at the very least, prior research that has proposed monitoring, checking, and proactive maintenance of intentions provide a more reasonable account of cue detection and interference effects and seem much more likely to be at play in both laboratory decision making tasks and naturalistic PM tasks that people complete daily. That being said, we agree with the proponents of delay theory that we should constantly be pushing the boundaries of how we think about the mechanisms underlying PM and hope that our work adds to that endeavor.

References

Anderson, F., & McDaniel, M. (2019). Anderson, F. T., & McDaniel, M. A. (2019). Retrieval in prospective memory: Multiple processes or just delay? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 72,* 2197–2207. http://dx.doi.org/10.1177/1747021819845622

Anderson, F. T., McDaniel, M. A., & Einstein, G. O. (2017). Remembering to remember: An examination of the cognitive processes underlying prospective memory. *Learning and memory: A comprehensive reference, 2,* 451–463.

Anderson, F. T., Rummel, J., & McDaniel, M. A. (2018). Proceeding with care for successful prospective memory: Do we delay ongoing responding or actively monitor for cues? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44,* 1036.

Anderson, F. T., Strube, M. J., & McDaniel, M. A. (2019). Toward a better understanding of costs in prospective memory: A meta-analytic review. *Psychological Bulletin, 145,* 1053–1081. http://dx.doi.org/10.1037/bul0000208

Ball, B. H. (2015). Cognitive control processes underlying continuous and transient monitoring processes in event-based prospective memory (Doctoral dissertation). Arizona State University, Tucson, AZ.

Ball, B. H., & Aschenbrenner, A. J. (2018). The importance of age-related differences in prospective memory: Evidence from diffusion model analyses. *Psychonomic Bulletin & Review, 25,* 1114–1122. http://dx.doi.org/10.3758/s13423-017-1318-4

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,* 793–811.

Ball, B. H., Brewer, G. A., Loft, S., & Bowden, V. (2015). Uncovering continuous and transient monitoring profiles in event-based prospective memory. *Psychonomic Bulletin & Review, 22,* 492–499. http://dx.doi.org/10.3758/s13423-014-0700-8

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

Ball, B. H., Vogel, A. E., & Brewer, G. A. (2019). Individual differences in prospective memory. In J. Rummel & M. McDaniel (Eds.), *Current issues in memory: Prospective memory* (pp. 116–134). London, UK: Taylor Francis. http://dx.doi.org/10.4324/9781351000154

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459. http://dx.doi.org/10.3758/BF03193014

Boag, R. J., Strickland, L., Heathcote, A., Neal, A., & Loft, S. (2019). Cognitive control and capacity for prospective memory in complex dynamic environments. *Journal of experimental psychology. General.* http://dx.doi.org/10.1037/xge0000599

Boag, R. J., Strickland, L., Loft, S., & Heathcote, A. (2019). Strategic attention and decision control support prospective memory in a complex dual-task environment. *Cognition, 191,* 103974. http://dx.doi.org/10.1016/j.cognition.2019.05.011

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). Fixed-effect versus random-effects models. *Introduction to Meta-analysis, 77,* 85.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences, 16,* 106–113. http://dx.doi.org/10.1016/j.tics.2011.12.010

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,* 304–311. http://dx.doi.org/10.3758/MC.38.3.304

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57,* 153–178. http://dx.doi.org/10.1016/j.cogpsych.2007.12.002

Bugg, J. M., McDaniel, M. A., & Einstein, G. O. (2013). Event-based prospective remembering: An integration of prospective memory and cognitive control theories. In D. Reisberg (Ed.), *The Oxford handbook of cognitive psychology* (pp. 267–282). New York, NY: Oxford University Press.

Bugg, J. M., McDaniel, M., Scullin, M. K., & Braver, T. S. (2011). Revealing list-level control in the Stroop task by uncovering its benefits and a cost. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1595–1606. http://dx.doi.org/10.1037/a0024670

Burgess, P. W., Quayle, A., & Frith, C. D. (2001). Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia, 39,* 545–555. http://dx.doi.org/10.1016/S0028-3932(00)00149-4

Cona, G., Bisiacchi, P. S., Sartori, G., & Scarpazza, C. (2016). Effects of cue focality on the neural mechanisms of prospective memory: A meta-analysis of neuroimaging studies. *Scientific Reports, 6,* 25983. http://dx.doi.org/10.1038/srep25983

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp. 409–422). Amsterdam, the Netherlands: Elsevier.

DeWitt, M. R., Knight, J. B., Hicks, J. L., & Ball, B. H. (2012). The effects of prior knowledge on the encoding of episodic contextual details. *Psychonomic Bulletin & Review, 19,* 251–257. http://dx.doi.org/10.3758/s13423-011-0196-4

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science, 14,* 286–290. http://dx.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,* 147–162. http://dx.doi.org/10.1037/a0019184

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,* 327–342. http://dx.doi.org/10.1037/0096-3445.134.3.327

Einstein, G. O., McDaniel, M. A., Williford, C. L., Pagan, J. L., & Dismukes, R. (2003). Forgetting of intention demands is rapid. *Journal of Experimental Psychology: Applied, 9,* 147–162. http://dx.doi.org/10.1037/1076-898X.9.3.147

Gordon, B. A., Shelton, J. T., Bugg, J. M., McDaniel, M. A., & Head, D. (2011). Structural correlates of prospective memory. *Neuropsychologia, 49,* 3795–3800. http://dx.doi.org/10.1016/j.neuropsychologia.2011.09.035

Guynn, M. J. (2003). A two-process model of strategic monitoring in event-based prospective memory: Activation/retrieval mode and checking. *International Journal of Psychology, 38,* 245–256. http://dx.doi.org/10.1080/00207590344000178

Guynn, M. J., McDaniel, M. A., & Einstein, G. O. (2001). Remembering to perform intentions: A different type of memory? In H. D. Zimmer, R. L. Cohen, M. J. Guynn, J. Engelkamp, R. Kormi-Nouri, & M. A. Foley (Eds), *Memory for action: A distinct form of episodic memory?* (pp. 25–48). Oxford, UK: Oxford University Press.

Hak, T., Van Rhee, H. J., & Suurmond, R. (2016). *How to interpret results of meta-analysis (version 1.0)*. Rotterdam, the Netherlands: Erasmus Rotterdam Institute of Management. Retrieved from https://ssrn.com/abstract=3241367

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

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

Hicks, J. L., Marsh, R. L., & Cook, G. I. (2005). Task interference in time-based, event-based, and dual intention prospective memory conditions. *Journal of Memory and Language, 53,* 430–444. http://dx.doi.org/10.1016/j.jml.2005.04.001

Horn, S. S., & Bayen, U. J. (2015). Modeling criterion shifts and target checking in prospective memory monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 95–117. http://dx.doi.org/10.1037/a0037676

Loft, S., Bowden, V. K., Ball, B. H., & Brewer, G. A. (2014). Fitting an ex-Gaussian function to examine costs in event-based prospective memory: Evidence for a continuous monitoring profile. *Acta Psychologica, 152,* 177–182. http://dx.doi.org/10.1016/j.actpsy.2014.08.010

Loft, S., & Remington, R. W. (2013). Brief delays in responding reduce focality effects in event-based prospective memory. *The Quarterly Journal of Experimental Psychology, 66,* 1432–1447. http://dx.doi.org/10.1080/17470218.2012.750677

McDaniels, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval. *Applied Cognitive Psychology, 14,* S127–S144. http://dx.doi.org/10.1002/acp.775

McDaniel, M. A., LaMontagne, P., Beck, S. M., Scullin, M. K., & Braver, T. S. (2013). Dissociable neural routes to successful prospective memory. *Psychological Science, 24,* 1791–1800. http://dx.doi.org/10.1177/0956797613481233

McDaniel, M. A., Umanath, S., Einstein, G. O., & Waldum, E. R. (2015). Dual pathways to prospective remembering. *Frontiers in Human Neuroscience, 9,* 392. http://dx.doi.org/10.3389/fnhum.2015.00392

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108. http://dx.doi.org/10.1037/0033-295X.85.2.59

Redick, T. S., & Engle, R. W. (2011). Rapid communication: Integrating working memory capacity and context-processing views of cognitive control. *Quarterly Journal of Experimental Psychology, 64,* 1048–1055.

Richmond, L. L., Redick, T. S., & Braver, T. S. (2015). Remembering to prepare: The benefits (and costs) of high working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1764–1777. http://dx.doi.org/10.1037/xlm0000122

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-prime (version 2.0) [Computer software and manual]. Pittsburgh, PA: Psychology Software Tools Inc.

Schnitzspahn, K. M., Stahl, C., Zeintl, M., Kaller, C. P., & Kliegel, M. (2013). The role of shifting, updating, and inhibition in prospective memory performance in young and older adults. *Developmental Psychology, 49,* 1544–1553. http://dx.doi.org/10.1037/a0030579

Scullin, M. K., McDaniel, M. A., & Shelton, J. T. (2013). The dynamic multiprocess framework: Evidence from prospective memory with contextual variability. *Cognitive Psychology, 67,* 55–71. http://dx.doi.org/10.1016/j.cogpsych.2013.07.001

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,* 736–749. http://dx.doi.org/10.1037/a0018971

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Shelton, J. T., & Scullin, M. K. (2017). The dynamic interplay between bottom-up and top-down processes supporting prospective remembering. *Current Directions in Psychological Science, 26,* 352–358. http://dx.doi.org/10.1177/0963721417700504

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,* 347.

Smith, R. E. (2010). What costs do reveal and moving beyond the cost debate: Reply to Einstein and 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,* 347–361. http://dx.doi.org/10.1037/0278-7393.29.3.347

Smith, R. E., Hunt, R. R., McVay, J. C., & McConnell, M. D. (2007). The cost of event-based prospective memory: Salient target events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 734–746. http://dx.doi.org/10.1037/0278-7393.33.4.734

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., . . . Tetzlaff, J. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal (Clinical Research Ed.), 343,* d4002. http://dx.doi.org/10.1136/bmj.d4002

Strickland, L., Elliott, D., Wilson, M. D., Loft, S., Neal, A., & Heathcote, A. (2019). Prospective memory in the red zone: Cognitive control and capacity sharing in a complex, multi-stimulus task. *Journal of Experimental Psychology: Applied, 25,* 695–715.

Strickland, L., Heathcote, A., Remington, R. W., & Loft, S. (2017). Accumulating evidence about what prospective memory costs actually reveal. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43,* 1616–1629. http://dx.doi.org/10.1037/xlm0000400

Strickland, L., Loft, S., Remington, R. W., & Heathcote, A. (2018). Racing to remember: A theory of decision control in event-based prospective memory. *Psychological Review, 125,* 851–887. http://dx.doi.org/10.1037/rev0000113

Zuber, S., Kliegel, M., & Ihle, A. (2016). An individual difference perspective on focal versus nonfocal prospective memory. *Memory & Cognition, 44,* 1192–1203. http://dx.doi.org/10.3758/s13421-016-0628-5

Received June 19, 2019
Revision received August 17, 2020
Accepted September 11, 2020 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.