Journal of Experimental Psychology:
Learning, Memory, and Cognition
2010, Vol. 36, No. 3, 736–749

0278-7393/10/$12.00 DOI: 10.1037/a0018971

Focal/Nonfocal Cue Effects in Prospective Memory: Monitoring Difficulty or Different Retrieval Processes?

Michael K. Scullin, Mark A. McDaniel, Jill T. Shelton, and Ji Hae Lee
Washington University

We investigated whether focal/nonfocal effects (e.g., Einstein et al., 2005) in prospective memory (PM) are explained by cue differences in monitoring difficulty. In Experiment 1, we show that syllable cues (used in Einstein et al., 2005) are more difficult to monitor for than are word cues; however, initial-letter cues (in words) are similar in monitoring difficulty to word cues (Experiments 2a and 2b). Accordingly, in Experiments 3 and 4, we designated either an initial letter or a particular word as a PM cue in the context of a lexical decision task, a task that presumably directs attention to focal processing of words but not initial letters. We found that the nonfocal condition was more likely than the focal condition to produce costs to the lexical decision task (task interference). Furthermore, when task interference was minimal or absent, focal PM performance remained relatively high, whereas nonfocal PM performance was near floor (Experiment 4). Collectively, these results suggest that qualitatively different retrieval processes can support prospective remembering for focal versus nonfocal cues.

Keywords: prospective memory, spontaneous retrieval, strategic monitoring, cue focality, intentions

Prospective memory (PM) refers to remembering to perform an intended action in the future. Common PM intentions include taking medication, attaching files to e-mails, and showing up for appointments. One focus of PM research has been determining the processes that underlie remembering to perform such intentions. The preparatory attentional and memory processes (PAM) theory (Smith, 2003; Smith & Bayen, 2004; Smith, Hunt, McVay, & McConnell, 2007) argues that an intention may be retrieved only if the individual engages a preparatory process such as monitoring the environment for cues that signal that the intention should be executed. Because monitoring (and other preparatory) processes are strategic and nonautomatic, it is assumed that engaging and maintaining these processes requires attentional resources that would otherwise be devoted to performing ongoing activities. Therefore, according to PAM theory, delaying performance of a PM intention should always interfere with performance on ongoing activities (to the extent that prospective remembering is evident). Evidence for such task interference has often been demonstrated as slower performance of the ongoing task when a PM task is embedded relative to when the ongoing task is performed alone.

Task interference has been consistently demonstrated in certain situations (e.g., when 6 PM target cues are used; Cohen, Jaudas, & Gollwitzer, 2008; Smith, 2003; and when the target item is not focal to the ongoing task; Einstein et al., 2005; McDaniel, Einstein, & Rendell, 2008). However, not all research has reported significant task interference, even though performance on the PM task has remained relatively high (e.g., Cohen et al., 2008; Einstein et al., 2005). Such reports of nonsignificant task interference are critical because the finding that prospective remembering occurs in the absence of monitoring suggests that an additional retrieval process exists.

Consistent with the above patterns, the multiprocess theory suggests that qualitatively different retrieval processes can be engaged to support prospective remembering (Einstein et al., 2005; McDaniel & Einstein, 2000; McDaniel & Einstein, 2007a). One such retrieval process, referred to as (strategic) monitoring, is akin to the attention-demanding monitoring process posited by the PAM theory. Another set of processes allows spontaneous retrieval of an intention without engaging resource-demanding monitoring processes (for details, see McDaniel & Einstein, 2007a; McDaniel, Guynn, Einstein, & Breneiser, 2004). The term *spontaneous retrieval* reflects the assumption that though some aspects of retrieval may be automatic (such as noticing that a cue is special in some way; Marsh, Hicks, & Watson, 2002), prospective remembering is unlikely to be fully automatic (Einstein & McDaniel, in press; McDaniel & Scullin, 2010). For example, according to the discrepancy plus attribution view, detecting that a PM cue is discrepant is relatively automatic, but attributing that discrepancy to the PM intention may be effortful (Breneiser & McDaniel, 2006). Likewise, the reflexive associative view is that focally processing a PM cue may reflexively trigger retrieval of that intention but that resources are necessary to keep the retrieved intention in mind and coordinate its execution with that of the ongoing task (McDaniel, Robinson-Riegler, & Einstein, 1998; McDaniel & Scullin, 2010; also cf. Marsh et al., 2002). Therefore, a spontaneous retrieval process does not confer automaticity in PM performance, but instead supports prospective remembering when

Michael K. Scullin, Mark A. McDaniel, Jill T. Shelton, and Ji Hae Lee, Department of Psychology, Washington University.

While conducting this research, Jill T. Shelton was supported by National Institute on Aging Grant T32 AG000030-32. We thank Kristy Duprey and Michael Clerkin for their help collecting data and Christopher Kurby and Nathan Rose for their helpful comments in preparing this article.

Correspondence concerning this article should be addressed to Michael K. Scullin, Department of Psychology, Washington University, St. Louis, MO 63130-4899. E-mail: mscullin@wustl.edu

monitoring is not engaged (for further elaboration on the mechanisms that stimulate spontaneous retrieval processes, see McDaniel & Einstein, 2007b; McDaniel et al., 2004).

For present purposes, one factor that has been identified as potentially influencing the type of process (i.e., monitoring or spontaneous retrieval) required to support PM retrieval is the conjoint nature of the ongoing and PM tasks. In some cases, the encoded features of the PM cue are processed in the service of the ongoing task (i.e., ongoing task processing stimulates extraction of the originally encoded features of the PM cue). For example, if a particular word is designated as the PM cue, then presumably the normative semantic features of the word are activated and associated with the intention during encoding. If the subsequent ongoing task requires processing of the semantic features that were activated and associated with the intention during encoding, then the PM cue can be considered focal to the ongoing task. McDaniel and colleagues have suggested that focal PM cues stimulate spontaneous retrieval of the associated intention (Einstein & McDaniel, 2005; Einstein et al., 2005; McDaniel & Einstein, 2000, 2007a).

In other cases, the ongoing task does not direct attention toward processing the features of the PM cue. For example, a particular syllable might be designated as the PM cue, but the ongoing task may require semantic processing of words (not syllabic processing). In this case, the PM cue can be considered nonfocal. According to the multiprocess theory, because the pertinent features of nonfocal PM cues are not normally processed during the ongoing task, successful nonfocal PM performance depends upon the engagement of attention-demanding processes (e.g., monitoring) directed toward identifying the PM cue. Therefore, the multiprocess theory predicts that task interference (reflective of monitoring) is likely to emerge and will be associated with PM performance when a nonfocal PM cue is used. By contrast, focal PM cues can trigger PM retrieval even when resource-demanding monitoring processes are not present (i.e., task interference is not observed).

Einstein et al. (2005, Experiments 1 and 2) investigated task interference for focal and nonfocal cues by varying whether a word (focal) or syllable (nonfocal) PM target cue was embedded within a category judgment task. In two experiments, significant task interference was demonstrated in the nonfocal but not the focal condition. Critically, PM task performance was high in the focal condition despite evidence that participants were not monitoring for the PM cues. Apparently with focal PM cues, retrieval of the PM intention is relatively spontaneous, thereby supporting good PM performance with minimal or no interference to the ongoing task.

Though suggestive, the above pattern does not decisively adjudicate between the PAM theory and the multiprocess theory. The PAM theory might accommodate the Einstein et al. (2005) pattern by suggesting that monitoring for words is easier than monitoring for syllables, independent of the PM task demands. If this were the case, then task interference in the focal condition (word cue) would be expected to be significantly dampened relative to task interference observed in the nonfocal condition (syllable cue). Accordingly, we first conducted a speed of cue detection experiment (not a PM experiment) to test the possibility that it is more difficult to monitor for syllables than for words (Experiment 1). We investigated this issue by asking participants to respond as quickly as possible whether a presented word was a target word (or not) in some blocks of trials or whether a presented word contained a target syllable (or not) in other blocks. To foreshadow, the results of Experiment 1 suggested that it is easier to monitor for a word than a syllable. Note that this finding would not necessarily rule out the multiprocess theory's interpretation of the previously reported differences in performance patterns between focal and nonfocal cues (Einstein et al., 2005), but the finding introduces ambiguity into the theoretical interpretation of such results.

To eliminate such ambiguity, we subsequently attempted to identify a pair of cue types that did not differ in monitoring difficulty (Experiments 2a and 2b). Experiments 3 and 4 then used the cue types that were shown in Experiments 2a and 2b to be equivalent on monitoring difficulty as PM targets in an ongoing task that directed attention to processing the relevant feature of one cue type but not the other. By doing so, we were able to more convincingly test the multiprocess theory's interpretation of divergent effects of focal/nonfocal PM cues (suggesting qualitatively different retrieval processes).

Experiment 1

We first tested whether Einstein et al.'s (2005) focal/nonfocal effects could have been related to differences in cue-monitoring difficulty. To do so, we gave participants short blocks of a cue-monitoring task in which they responded as quickly as possible about whether a presented word was a target (word or syllable) or not. Because our goal was to observe the speed of the monitoring process per se, the target trials were presented relatively frequently (15% of trials) so as to encourage monitoring on every trial (see Loft & Yeo, 2007, for evidence that a similar target frequency encourages monitoring). We reasoned that if speed of responding is slower during syllable blocks than word blocks, then Einstein et al.'s observed differences in task interference across the nonfocal (syllable target) and focal (word target) conditions might be explained by differences in cue-monitoring difficulty (rather than their assumption that different retrieval processes were associated with nonfocal and focal PM cues).

Method

Participants and design. Twenty-four Washington University undergraduates participated for partial class credit or $5 monetary compensation. Participants were tested in groups of 1–4 during sessions that lasted approximately 20 min. The type of cue-monitoring task (syllable or word) varied within subjects.

Procedure. Participants first learned the instructions for the cue-monitoring task. They were told that they would see items on the computer screen and would be asked to respond as quickly as possible whether the item was a target item or not a target item by pressing the keys labeled *Y* and *N* (1 and 2 on the number pad, respectively). Further, participants were instructed that they should aim for 95%–100% accuracy while responding as quickly as they could.

Participants monitored for syllable targets first and word targets second or vice versa. Participants who performed the syllable monitoring task first were told they would be searching for a target syllable that changed between blocks. They were then given a 10-trial practice block in which the practice target syllable *zen* (or target word *zenith* in counterbalance) appeared three times. Following each practice trial, participants received feedback regarding their accuracy and speed of responding. After the practice block, participants were given a chance to ask questions before beginning the experimental blocks. Each of the 10 experimental blocks included 10 filler (i.e., nontarget) trials and 1–2 target trials. Participants were given a new target syllable before each block. The order of target blocks and filler trials within blocks was random for each participant. During the five critical blocks, the target syllable occurred only once on a randomly determined trial in the latter half of the block. During the other five blocks, the target syllable appeared twice in randomized locations, once during the first five trials in the corresponding target word, and again during the last five trials in a different word. We included some blocks with two target trials to encourage continuous monitoring. The target trial used during the one-target trial block was equivalent across the word and the syllable conditions (and initial-letter condition in Experiments 2a and 2b); thus, all statistical analyses focused on comparing target types during these critical blocks (blocks with one target trial).

Participants who monitored for syllable targets during the first half of the experiment were told that they would search for target words next. They then performed a block of 10 practice trials during which their target word *bishop* (or target syllable *bish* in a counterbalanced condition) appeared three times. Speed and accuracy feedback were given after each practice trial. Following the practice trials, participants were allowed to ask questions before performing the experimental blocks. The second half of experimental blocks was identical in structure to the first half; participants performed 10 randomly ordered blocks (new target word in each block) consisting of 10 filler trials and one or two target trials each. After completing the second half of the cue-monitoring task, participants were debriefed, thanked, and excused.

Materials. Two lists (order counterbalanced) of lowercase target and filler words were generated from the Balota et al. (2007) English lexicon database. All words were 6–12 letters and 2–4 syllables in length and had a log-transformed Hyperspace Analogue to Language (HAL) frequency within one standard deviation of the mean (i.e., 3.76–8.56). The targets for Experiment 1 (as well as for Experiments 2a and 2b) are listed in Appendix A.

Results and Discussion

Unless otherwise reported, all statistical tests were conducted with a significance level of .05. Estimates of effect size (η_p^2) are reported for all F values greater than 1 so as to facilitate comparison of effects within the article.

We first examined whether identifying target syllables was more difficult than identifying target words by examining mean response accuracy (i.e., proportion correct) on target and filler trials. An analysis of variance (ANOVA) that included target type (word or syllable) as a within-subjects variable revealed no effect of target type for either target trial accuracy (M_{word} = .89 and $M_{syllable}$ = .86; $F < 1$) or filler trial accuracy, $F(1, 23) = 1.53$, $MSE = .489$, $\eta_p^2 = .06$ (see Table 1).

The more critical question is whether there was a monitoring speed difference between target conditions. Following Einstein et al.'s (2005) trimming method, we examined mean reaction times (RTs) on correct filler trials (i.e., "no" responses) that were no greater than two standard deviations from each individual's mean (syllable and word blocks were trimmed separately). A within-subjects ANOVA that included target type (syllable or word) confirmed that mean trimmed filler RTs were reliably greater in the syllable condition than in the word condition, $F(1, 23) = 93.11$, $MSE = 4{,}024.37$, $\eta_p^2 = .80$ (see Table 1).

Table 1
Mean Filler Trial Accuracy (Proportion Correct) and Trimmed Reaction Times (RTs; in Milliseconds) on the Cue-Monitoring Task in Experiments 1 and 2a

Variable	Word	Syllable	Initial letter
	Experiment 1		
Accuracy	.998 (.01)	.993 (.02)	
RT	372 (61)	549 (121)	
	Experiment 2a		
Accuracy	.998 (.01)		.997 (.01)
RT	376 (43)		385 (47)

Note. Standard deviations are in parentheses.

The results of the present experiment indicated that monitoring for syllables was more difficult than monitoring for words. Importantly, these results suggest an alternative to the multiprocess theory's interpretation that word and syllable PM cues produced different levels of task interference (and PM task performance) in the Einstein et al. (2005) research because the word cue (focal) stimulated spontaneous retrieval whereas the syllable cue (nonfocal) required attention-demanding monitoring processes to support prospective remembering. The current results allow the alternative possibility that prospective remembering required monitoring regardless of the cue condition, with the differences between the word and syllable-cue conditions simply reflecting the differential difficulty of monitoring for these cues.

Experiment 2a

One approach to legislating between the alternative theoretical interpretations of patterns of task interference produced by focal and nonfocal cues would be to identify conditions in which the difficulty of monitoring for focal and nonfocal cues were approximately equivalent. As the first step toward implementing this experimental strategy, we examined a type of cue (words that began with a target letter, referred to as *initial-letter* cues) that we believed would be no more difficult to monitor for than word cues. We repeated Experiment 1 but contrasted the initial-letter cue with a word cue.

It is important to note that extant monitoring theories (as well as the multiprocess theory) are silent regarding the dynamics by which a putative monitoring process might be applied on a trial. In the PAM model, the monitoring process initiates a check to evaluate whether the presented stimulus is the PM cue, but it remains unclear whether such a check is presumed to occur before or after processing for the ongoing task is initiated (see Smith, 2003). In the absence of existing theory, we assumed that either could occur depending on individual differences, task demands, and other contextual factors. Accordingly, we examined monitoring difficulty both under the assumption that the monitoring process would be applied immediately upon stimulus presentation (Experiment 2a) as well as the assumption that there may be individual differences in whether the monitoring process would occur before or following ongoing task processing (Experiment 2b). We reasoned that if monitoring performances were equivalent across the target cue types (word, initial-letter), then we would have identified cues that were equivalent in terms of monitoring difficulty, regardless of when monitoring was recruited (i.e., before or after ongoing task processing). As in Experiment 1, targets constituted 15% of trials in the present experiment so as to encourage monitoring on every trial and provide a comparison of the speed of monitoring (per se) between word cues and initial-letter cues.

Method

The participants were 37 Washington University undergraduates who received partial class credit or $5 monetary compensation for their time. The procedure was identical to Experiment 1 except that the syllable condition was replaced by an initial-letter condition in which participants searched for items that began with a target letter. The materials were identical to Experiment 1, and the targets are specified in Appendix A.

Results and Discussion

All analyses were conducted in the same manner as in Experiment 1. We first examined whether accuracy on filler and target trials differed between the word and initial-letter conditions by conducting separate within-subjects (word or initial-letter target type) ANOVAs for filler trials and for target trials. Performance accuracy did not differ between target conditions on target trials (both Ms = .83; $F < 1$) or filler trials, $F(1, 36) = 1.30$, $MSE = .09$, $\eta_p^2 = .03$ (see Table 1).

We next examined whether monitoring was more difficult for target initial letters than target words by comparing mean trimmed RTs on filler items (see Table 1). Consistent with the previous accuracy analyses, there was no effect of target type on mean trimmed RT, $F(1, 36) = 2.71$, $MSE = 601.06$, $\eta_p^2 = .08$ (.97 power to detect a medium-sized effect).[1] Thus, the results of the present experiment suggested that monitoring for a particular initial letter is similar in difficulty to monitoring for a particular word.

Experiment 2b

The results of Experiment 2a demonstrated that monitoring difficulty was similar for word and initial-letter cues in a situation in which cue monitoring was assumed to occur prior to any other ongoing task decision (i.e., the primary response in Experiment 2a was to monitor for a target). However, Experiment 2a may have differed in important ways from typical PM experiments. Accordingly, in the following experiment we made two important changes to our evaluation of cue-monitoring difficulty. First, to more closely parallel the target frequencies present during a PM task, we reduced the frequency of target trials from 15% to 5% of all trials (target frequency is below 10% in most PM studies). Second, instead of requiring participants to monitor immediately on each trial (Experiment 2a), we encouraged participants to monitor in the manner that felt most natural to them (i.e., either before or after making a lexical decision). This allowed us to gain insight into the typical monitoring strategy employed as well as closely match the typical PM task situation in which participants could monitor before or after ongoing task processing.

Method

Participants. The 32 participants received $5 monetary compensation for their time.

Procedure. The major difference between the current procedure and that used in Experiment 2a was that participants were required to make both a lexical decision and a target decision on each trial. Participants were first told that they would perform two tasks in the experiment, the lexical decision task (referred to as the *SPEED* task) and the cue-monitoring task. For the lexical decision task, they were instructed to indicate whether a string of letters formed a real word or not by pressing the keys labeled *Word YES* and *Word NO* (the *A* and *S* keys, respectively). The importance of performing the lexical decision task as quickly and as accurately as possible was emphasized. Following the instructions, participants practiced the lexical decision task for 20 trials and received feedback regarding the speed and accuracy of their responses after each trial. After a lexical decision practice block, participants received the cue-monitoring task instructions and cue-monitoring practice block, which were identical to the previous experiments except that the response keys were labeled *Target YES* and *Target NO* (the *K* and *L* keys, respectively). Participants were also told (emphasized by presenting this information in capital letters) that the targets would only occur on word trials. Both instruction order (i.e., lexical decision task first or cue-monitoring task first) and key set (*A* and *S*, or *K* and *L*) were counterbalanced across participants.

After participants practiced both the lexical decision task and the cue-monitoring task alone, we told them that they would always make both word/nonword decisions and target decisions on the same trial. They were instructed to not attempt to press the word and target keys simultaneously. Participants were told that it did not matter whether they made a word or target decision first, and to respond in whichever order felt most natural to them. After these instructions, participants were given a 20-trial practice block in which their practice target word *zenith* (or words that began in the letter *z*) occurred three times. During the practice block, and later during the experimental blocks, the word (or nonword) remained on the screen until two responses were recorded and a blank screen appeared for 500 ms between trials. The experimental blocks were structured in the same manner as in the previous experiments (e.g., monitor for word and initial-letter targets in 10 blocks each, with a new target for each block) except that we extended the length of each block from 10 trials to 30 trials by adding more filler trials to each block (the number of targets remained the same). We increased the number of filler trials in

[1] Power analyses were conducted using G*Power statistical software (Faul, Erdfelder, Lang, & Buchner, 2007). There was no a priori reason to expect a certain effect size for cue-monitoring differences (except perhaps the large effect obtained in Experiment 1) and so we report power to detect a medium-sized effect (.25) in Experiments 2a and 2b. For Experiment 3, we relied on the ability to demonstrate task interference during some parts of the ongoing task to assure that we had ample power to detect significant task interference effects. In Experiment 4, we followed Scullin et al.'s (2010) procedure of reporting power to detect medium-sized (r = .25), medium- to small-sized (r = .15), and small-sized (r = .10) effects. The task interference literature suggests that, if present, focal task interference will be observed as a medium-sized effect (Smith et al., 2007).

each block to decrease the total frequency of target trials in the experiment.

Materials. The target items were identical to Experiment 2a (see Appendix A) and we counterbalanced the order in which participants monitored for target initial letters or target words in the same manner as in the previous experiments. For the filler items, two lists of 100 words were generated from the Balota et al. (2007) lexicon database according to the same specifications as in Experiment 1. Half of the words in each list were changed to nonwords by repositioning one or two letters. During the first 10 experimental blocks, filler items (i.e., words and nonwords that were not targets) were drawn randomly from one list, whereas during the last 10 experimental blocks filler items were randomly drawn from the second list (list order counterbalanced).

Results and Discussion

Consistent with the previous experiments, analyses were conducted on critical blocks.

Monitoring strategy. We first determined whether there was a bias for monitoring before or after lexical decision processing. We calculated the proportion of trials in which a lexical decision response was recorded first and conducted a $2 \times 2 \times 2 \times 2$ mixed ANOVA that included target type (word or initial letter) and trial type (filler or target) as within-subjects factors as well as instruction order (lexical decision instructions first or second) and key order (lexical decisions with left or right hand) as between-subjects factors. There were no significant effects (largest $F[1, 28] = 3.96$, $\eta_p^2 = .12$, for trial type main effect). Interestingly, the first response on each trial was just as likely to be a lexical decision ($M = 0.48$) as a target decision ($M = 0.52$). Further investigation revealed that the majority of individuals adopted a consistent response strategy (same response order used on at least 98% of the trials); there were a similar number of participants who (almost) always made a lexical decision first ($n = 15$) compared to individuals who (almost) always made a target decision first ($n = 14$).

Cue-monitoring difficulty. Target detection accuracy was first examined as target decision accuracy on target trials. A target decision was counted as correct if the *Target YES* key was pressed on the target trial (unless the *Target NO* key was initially pressed, as occurred on 3.4% of target trials). Using a more lenient criterion (i.e., target trials in which the *Target YES* key was pressed following an initial *Target NO* keypress were included as correct) did not alter the results. A 2×2 mixed ANOVA that included target type (word or initial letter) as a within-subjects variable and monitoring strategy (lexical decision first or target decision first) as a between-subjects variable demonstrated that individuals ($n = 15$) who typically made lexical decisions first ($M = 0.81$) had greater target decision accuracy on target trials than did participants ($n = 14$) who made target decisions first ($M = 0.62$), $F(1, 27) = 6.62$, $MSE = .03$, $\eta_p^2 = .20$. However, monitoring strategy did not interact with target type ($F < 1$), and importantly, target decision accuracy did not significantly differ across target type ($M_{word} = .72$ and $M_{letter} = .71$; $F < 1$). These similar accuracy levels across word and initial-letter blocks held when all 32 participants (not just those classified as monitor first or second) were included in a within-subjects ANOVA ($F < 1$). Furthermore, lexical decision accuracy on target trials also did not differ during word ($M = 0.93$) and initial-letter ($M = 0.89$) blocks ($F < 1$).

We next examined target decision and lexical decision accuracy on filler trials by conducting a 2×2 within-subjects ANOVA that included target type (word or initial letter) and decision type (lexical decision or target decision) as independent variables (monitoring strategy had no effect on filler trial accuracy and was therefore not included in this analysis). There was a decision type main effect, $F(1, 31) = 179.51$, $MSE = .0005$, $\eta_p^2 = .85$, such that target decisions ($M = 0.99$) were more accurate than lexical decisions ($M = 0.90$). There was neither a target type main effect or a Target Type $\times$ Decision Type interaction (both Fs < 1), demonstrating that filler trial accuracy was similar in the word and initial-letter conditions for lexical decisions ($M_{word} = .90$ and $M_{letter} = .91$) as well as target decisions (both Ms $> .99$).

Consistent with the previous experiments, we were most interested in comparing the target types on mean trimmed filler trial RTs. A within-subjects 2×2 ANOVA that included target type (word or initial letter) and decision type (lexical decision or target decision) produced no significant effects (all Fs < 1).[2] Importantly, despite adequate power to detect a medium-sized effect (.87), responding was similar during word blocks ($M = 1{,}001$ ms) compared to initial-letter blocks ($M = 1{,}017$ ms; $F < 1$). Thus, even when we compared word and initial-letter targets under conditions highly similar to those implemented in a typical PM task (i.e., low frequency of target trials and no constraints on monitoring strategy), these target types appear to be quite similar in monitoring difficulty. Therefore, our next goal was to conduct a PM experiment in which we embedded these target types in an ongoing task that encouraged focal processing of one target type but not the other.

Experiment 3

In this experiment, we gave participants the PM task of remembering to press a given key if they encountered a specified target cue (for some participants, the initial-letter *g*, and for other participants, the word *generous*) in the context of a lexical decision task. Collapsing across Experiments 2a and 2b, we found that monitoring for the initial-letter *g* was no more difficult than monitoring for the word *generous* when comparing mean trimmed filler trial RTs, target detection accuracy, or filler trial accuracy (all Fs < 1). Thus, we reasoned that if monitoring difficulty is the primary factor in observing task interference differences across

[2] To determine whether monitoring strategy influenced mean trimmed filler trial RTs, we conducted a $2 \times 2 \times 2$ mixed ANOVA that included target type (word or initial letter) and decision type (lexical decision or target decision) as within-subjects variables as well as monitoring strategy (lexical decision first or target decision first) as a between-subjects variable. The only significant effect was the Monitoring Strategy $\times$ Decision Type interaction, $F(1, 27) = 415.53$, $MSE = 5{,}192.07$, $\eta_p^2 = .94$ (next largest $F[1, 27] = 3.79$, $\eta_p^2 = .12$, for the decision type main effect). The interaction obtained because lexical decision responding ($M = 191$ ms) was quicker than target decision responding ($M = 788$ ms) in individuals ($n = 14$) who mostly made target decisions before lexical decisions, $F(1, 27) = 480.51$, $MSE = 5{,}192.07$, $\eta_p^2 = .95$; however, the reverse pattern obtained in individuals ($n = 15$) who typically made lexical decisions ($M = 884$ ms) before target decisions ($M = 161$ ms), $F(1, 27) = 755.09$, $MSE = 5{,}192.07$, $\eta_p^2 = .97$. Thus, the second response was much quicker than the first response, regardless of whether a lexical decision or a target decision occurred first.

focal and nonfocal cue conditions, then task interference (and PM task performance) should be equivalent across the word and initial-letter target cue conditions.

Alternatively, the hypothesis derived from the multiprocess theory is that focal PM cues (e.g., a particular word presented in the context of a lexical decision task) can support prospective remembering with no monitoring, whereas the identification of nonfocal PM cues (e.g., initial letters in the context of a lexical decision task) requires monitoring (Einstein et al., 2005; McDaniel & Einstein, 2000, 2007a). Accordingly, we hypothesized that task interference should be observed in the initial-letter condition but may not be present in the word condition. Because it is still possible for participants to monitor in a focal condition (Smith et al., 2007), it is important to identify instances in which monitoring is absent and to test whether focal prospective remembering still occurs (Einstein & McDaniel, in press; Einstein et al., 2005; Scullin, McDaniel, & Einstein, 2010). By the multiprocess theory but not the PAM theory, we reasoned that PM performance could be relatively high even in the absence of task interference for the focal condition. Therefore, we took two approaches to analyzing task interference across conditions. First, we examined overall task interference collapsed across all ongoing trials (Einstein et al., 2005; Smith et al., 2007). In addition, because overall task interference might have emerged if monitoring is engaged on only some trials by only some participants (Einstein & McDaniel, in press), we also examined task interference by dividing task blocks into quartiles (i.e., the set of trials preceding each target event). If monitoring is required for focal PM retrieval, then task interference must be present in each quartile for PM performance to remain relatively high. However, if spontaneous retrieval processes can support focal prospective remembering, then PM performance would remain relatively high even during quartiles in which task interference is absent.

The idea to examine task interference across quartiles raises the interesting question of how task interference may change across time. Previous research has demonstrated that task interference declines across task blocks (Loft, Kearney, & Remington, 2008; McDaniel et al., 2008). We therefore expected task interference in the focal condition to be absent in the final quartile (i.e., Quartile 4), possibly because participants realize that monitoring is not necessary to perform the PM intention, after detecting the focal target cue several times. Other research suggests that task interference may emerge after, but not before, the first target cue (Scullin, 2009). It is not yet entirely clear why such a pattern sometimes obtains, but one possible explanation is that, after spontaneously retrieving the PM intention on the first target cue, participants remember that they should also be performing the PM task (and expect the target cue to appear again shortly), and then monitor for the target for some duration thereafter. Though the cause of fluctuations in task interference is not completely understood, conducting task interference analyses across quartiles allowed us to examine the regularity of monitoring across task blocks, to illuminate whether monitoring changes would be accompanied by changes in PM performance, and to evaluate whether overall task interference is an appropriate measure for concluding that monitoring is always required for focal PM retrieval (as assumed by the PAM, but not the multiprocess, theory).

Method

Participants and design. Forty-eight Washington University undergraduates participated for partial class credit or $5 monetary compensation. They were tested in groups of 1–4 individuals in sessions lasting approximately 25 min. The design was a 2×2 mixed factorial in which PM target type (word or initial letter) varied between participants and block of the lexical decision task (PM or control) varied within participants.

Procedure. Participants performed two blocks of a lexical decision task. During one of these blocks, they were also required to perform a PM task (order of control and PM blocks were counterbalanced, with $n = 24$ in each).

Participants first learned the instructions for the lexical decision task that were identical to Experiment 2b. Following the instructions, participants practiced the lexical decision task for 20 trials and received speed and accuracy feedback after each trial. Then some participants received the instructions for the PM task. They were told that there was a secondary interest in their ability to remember to perform an action in the future and if they saw a particular item (*generous* in the word condition; an item beginning with the letter *g* in the initial-letter condition) that they should remember to press the *Q* key. Participants were reminded that their primary goal was to respond quickly to the lexical decision task and were required to explain their PM and lexical decision task instructions to the experimenter before continuing.

Next, participants performed an experimental block that included 10 buffer trials followed by 208 experimental trials (four PM targets, 102 words, 102 nonwords). The target item occurred on Trials 51, 102, 153, and 204 (in the initial-letter condition the words were *generous*, *grooming*, *glancing*, and *galleries*, respectively). During control blocks, the control word *inherent* (matched on number of letters and syllables to *generous*) was presented on these trials.

Following completion of the first experimental block, participants who performed the PM block first were instructed that they were entering into a new phase of the experiment. These participants were told (in capital letters to draw attention) to please note that they no longer needed to remember to press the *Q* key and that their only concern in the upcoming block was to perform the lexical decision task (see Scullin, Einstein, & McDaniel, 2009, for evidence of intention deactivation following these instructions). The participants who performed the control block first were given their PM task instructions and were asked to explain them to the experimenter. Participants then performed a second experimental block that was identical in structure to the first block. Following completion of the second block, participants were debriefed, compensated, and thanked for their participation.

Materials. Lowercase words and nonwords were generated from the Balota et al. (2007) lexicon database and were randomly assigned to one of two lists. In each list, 34 filler items (i.e., words and nonwords that were not target or buffer items) appeared once, 34 appeared twice, and 34 appeared three times. All filler items were 4–9 letters and 1–3 syllables in length. No filler words were forwardly associated with the word *generous* (according to the Nelson, McEvoy, & Schreiber, 1998, free association norms; see Scullin et al., 2010, Experiment 1, for the effects of filler items with forward associations to target words). The particular order of these lists was counterbalanced across participants.

Results and Discussion

Preliminary analyses demonstrated no main effect of block order and no interactions with block type (PM or control) or PM target type (word or initial letter) on PM task performance or lexical decision performance. Therefore, there were no practice or fatigue effects, and we report all analyses below with block order collapsed.

Overall ongoing task performance. We first evaluated lexical decision filler trial accuracy (number of correct trials divided by total possible correct). Furthermore, to ensure that participants were not selectively monitoring on word trials and not on nonword trials, we calculated the means for these trial types separately and included lexical decision trial as a variable in all analyses. Accuracy was included in a 2 × 2 × 2 mixed ANOVA in which PM target type (word or initial letter) was a between-subjects variable and block (PM or control) and lexical decision trial (words or nonwords) were within-subjects variables. There were no significant effects (largest $F[1, 46] = 3.68$, $\eta_p^2 = .07$, for the Block × PM Target Type interaction) as lexical decision accuracy was high in the word and initial-letter conditions during both PM blocks (both *M*s = .94) and control blocks ($M_{letter} = .95$ and $M_{word} = .94$).

Overall lexical decision task performance was also assessed using mean trimmed RTs. Consistent with the previous experiments, responses were trimmed to only include correct filler trials that were fewer than two standard deviations from each individual mean (Einstein et al., 2005). Trimming was conducted separately for word and nonword trials within control and PM blocks. These data were analyzed with a 2 × 2 × 2 mixed ANOVA that included PM target type (word or initial letter) as a between-subjects variable and block (PM or control) and lexical decision trial (words or nonwords) as within-subjects variables. RTs were significantly faster on word trials ($M = 508$ ms) than nonword trials ($M = 539$ ms), $F(1, 46) = 51.91$, $MSE = 288.07$, $\eta_p^2 = .53$ (this pattern was consistent throughout Experiment 3). More importantly, there was a main effect of block, $F(1, 46) = 29.73$, $MSE = 288.07$, $\eta_p^2 = .39$, that was qualified by a significant Block × PM Target Type interaction, $F(1, 46) = 8.01$, $MSE = 288.07$, $\eta_p^2 = .15$ (all other *F*s < 1). As displayed in Table 2 (see last row), the interaction emerged because there was greater overall task interference in the initial-letter condition than in the word condition. Because Experiments 2a and 2b demonstrated that monitoring for an initial letter is similar in difficulty to monitoring for a word, the target type differences in task interference are in accord with the expectations of the multiprocess theory's focal/nonfocal hypothesis but not the monitoring difficulty hypothesis.

A related central issue was whether the overall task interference results support the multiprocess theory's hypothesis that the focal condition might be associated with no monitoring. Planned comparisons revealed that overall task interference was not only significant in the initial-letter condition, $F(1, 46) = 57.03$, $MSE = 288.07$, $\eta_p^2 = .55$, but also in the word condition, $F(1, 46) = 6.00$, $MSE = 288.07$, $\eta_p^2 = .12$. The finding of overall task interference even in the word condition suggests that there was enough monitoring (by enough participants or on enough trials) to produce significant slowing on the ongoing task. However, as previously noted, overall task interference may be observed even when monitoring is absent proximal to target events (Scullin et al., 2010) and, therefore, may potentially mask the importance of spontaneous retrieval in supporting prospective remembering for many participants and on many target events (Einstein & McDaniel, in press). Therefore, we next attempted a more detailed analysis of task interference oriented toward individual target events by examining lexical decision task performance across block quartiles.

Ongoing task performance across quartiles. We first divided blocks into quartiles (i.e., the 50 trials preceding each PM target or control event) and averaged trimmed RTs on correct trials in each quartile (trimming was done separately for words and nonwords in control and PM blocks for each quartile). The data are presented in Table 2. Using the dependent measure of mean trimmed RTs in Quartile 1, we found that a mixed ANOVA including PM target type (word or initial letter) as a between-subjects variable as well as block (PM or control) and lexical decision trial (words or nonwords) as within-subjects variables revealed an effect of lexical decision trial, $F(1, 46) = 31.31$, $MSE = 966.04$, $\eta_p^2 = .40$; an effect of block, $F(1, 46) = 10.61$, $MSE = 966.04$, $\eta_p^2 = .19$; and a Block × PM Target Type interaction, $F(1, 46) = 15.37$, $MSE = 966.04$, $\eta_p^2 = .25$ (no other effects were significant; next largest $F[1, 46] = 3.75$, $\eta_p^2 = .08$, for the three-way interaction). Planned comparisons demonstrated that there was significant task interference in the initial-letter condition, $F(1, 46) = 18.89$, $MSE = 966.04$, $\eta_p^2 = .29$; however, in the word condition, there was no task interference ($F < 1$), as responding was nominally faster during PM blocks ($M = 510$ ms) than during

Table 2
Mean Trimmed Reaction Times (in Milliseconds) on the Lexical Decision Task as a Function of the PM Target Type, Lexical Decision Trial, Block, and Quartile in Experiment 3

	Word cue				Initial-letter cue			
	Words		Nonwords		Words		Nonwords	
Quartile	PM	Control	PM	Control	PM	Control	PM	Control
1	497 (64)	495 (54)	523 (70)	534 (83)	521 (55)	492 (90)	553 (81)	503 (62)
2	515 (64)	502 (60)	550 (86)	533 (81)	543 (64)	496 (57)	562 (84)	523 (60)
3	519 (71)	492 (61)	549 (68)	533 (75)	521 (76)	492 (55)	552 (74)	527 (64)
4	506 (55)	503 (71)	541 (71)	525 (74)	514 (66)	488 (61)	548 (74)	520 (57)
Overall	512 (61)	498 (57)	542 (70)	532 (77)	529 (59)	493 (61)	560 (74)	520 (59)

Note. Standard deviations are in parentheses. PM = prospective memory.

control blocks ($M = 514$ ms). Thus, in the word condition, there was no evidence for monitoring preceding the first target (word) cue.

Similar $2 \times 2 \times 2$ mixed ANOVAs were conducted for trimmed RTs during Quartiles 2, 3, and 4. In Quartile 2, there was a significant effect of lexical decision trial, $F(1, 46) = 25.74$, $MSE = 496.28$, $\eta_p^2 = .36$, block, $F(1, 46) = 29.93$, $MSE = 496.28$, $\eta_p^2 = .39$, and also a Block $\times$ PM Target Type interaction, $F(1, 46) = 7.04$, $MSE = 496.28$, $\eta_p^2 = .13$ (all other Fs < 1). Planned comparisons demonstrated that task interference was significant in both the initial-letter condition, $F(1, 46) = 46.81$, $MSE = 496.28$, $\eta_p^2 = .50$, and the word condition, $F(1, 46) = 5.44$, $MSE = 496.28$, $\eta_p^2 = .11$. The results of the Quartile 3 ANOVA also produced a main effect of lexical decision trial, $F(1, 46) = 43.42$, $MSE = 699.76$, $\eta_p^2 = .49$, and a main effect of block, $F(1, 46) = 11.98$, $MSE = 699.76$, $\eta_p^2 = .21$ (all other Fs < 1); the planned comparisons revealed significant task interference in the initial-letter condition, $F(1, 46) = 12.50$, $MSE = 699.76$, $\eta_p^2 = .21$, and in the word condition, $F(1, 46) = 7.56$, $MSE = 699.76$, $\eta_p^2 = .14$.

The Quartile 4 ANOVA demonstrated a main effect of lexical decision trial, $F(1, 46) = 29.76$, $MSE = 560.46$, $\eta_p^2 = .39$, as well as a main effect of block, $F(1, 46) = 5.79$, $MSE = 560.46$, $\eta_p^2 = .11$ (next largest $F[1, 46] = 1.29$, $\eta_p^2 = .03$, for the PM Target Type $\times$ Block interaction). However, planned comparisons showed that whereas task interference was significant in the initial-letter condition, $F(1, 46) = 10.36$, $MSE = 560.46$, $\eta_p^2 = .18$, task interference was not significant in the word condition, $F(1, 46) = 1.73$, $MSE = 560.46$, $\eta_p^2 = .04$. Thus, task interference was significant in each quartile in the initial-letter condition but only significant during the middle quartiles in the word condition. This pattern was fairly consistent with previous research (cf. patterns reported by Loft et al., 2008; McDaniel et al., 2008; Scullin, 2009) and provides good evidence that overall task interference may emerge because monitoring is present on some, but not all, ongoing task trials. Importantly, the finding in the word condition that task interference was present preceding the second and third target events but not the first and fourth target events allows a test of the prediction that prospective remembering does not depend on monitoring when a PM cue is focally processed (Einstein et al., 2005; Scullin et al., 2010). Under conditions of no monitoring (indexed by task interference), the PAM theory predicts PM performance to be at floor levels, whereas the multiprocess theory predicts that focal processing of a target cue will trigger spontaneous retrieval of the PM intention and thus allow for high levels of PM task performance.

PM task performance. We counted PM responses as correct if participants pressed the *Q* key on the target trial or on the following trial. To evaluate whether focal PM performance depended on monitoring processes being engaged in the quartile preceding the word target, we conducted a within-subjects ANOVA. Word PM performance was similar during the quartiles in which task interference was observed (Ms $= .83$ and $.88$ for Quartiles 2 and 3, respectively) relative to the quartiles in which task interference was absent (Ms $= .80$ and 1.00 for Quartiles 1 and 4, respectively; $F < 1$). These results are consistent with Scullin et al.'s (2010) finding that task interference observed on the five trials preceding focal PM targets was not associated with better PM performance relative to when task interference was absent. Furthermore, consistent with the multiprocess theory's prediction that spontaneous retrieval processes can support prospective remembering, PM performance was well above zero when task interference was absent (i.e., in Quartiles 1 and 4), $t(23) = 21.16$.

We next investigated whether there was a difference in PM task performance across conditions by conducting a between-subjects ANOVA that included PM target type (word or initial letter). Participants appeared to have remembered to press *Q* more often in the word condition ($M = 0.88$) than in the initial-letter condition ($M = 0.75$), but this difference was not significant, $F(1, 47) = 2.76$, $MSE = .07$, $\eta_p^2 = .06$. Such a nonsignificant result is not necessarily surprising because performance on the nonfocal task can be reasonably good when participants monitor (see, e.g., Smith & Bayen, 2004), and the task interference analyses suggested they were doing so in the initial-letter condition. Finally, a within-subjects ANOVA demonstrated that initial-letter PM performance did not change as a function of quartile ($\chi^2 < 1$; Ms $= .75, .75, .79$, and $.71$ for Quartiles 1, 2, 3, and 4, respectively). This result is consistent with the evidence that there was task interference in each quartile in the initial-letter condition.

The results of Experiment 3 may be summed as follows. First, embedding an initial-letter PM cue into an ongoing task that did not require processing of individual letters caused task interference. Task interference in this condition was observed before each initial-letter target cue, thereby suggesting that participants recruited resources to monitor for the cue so as to retrieve their intention at the appropriate moment. Second, embedding the word PM cue into the lexical decision task led to task interference when we averaged across all lexical decision trials (see Smith et al., 2007, for a similar finding). A more precise examination of task interference, however, revealed that task interference in the word condition was limited to the second and third quartiles. Third, even though there was no evidence for monitoring in the word condition before the first and fourth target cue, participants still remembered to perform the PM action on these trials 90% of the time. This last result was predicted only by the multiprocess theory and can be understood in terms of spontaneous retrieval processes supporting prospective remembering when a cue is focally processed.

Experiment 4

The results of Experiment 3 suggested that focal/nonfocal effects in PM tasks are not solely due to differences in monitoring difficulty. The purposes of Experiment 4 were to generalize this finding by using more than one target word/target letter and to test whether focal word and nonfocal initial-letter cues elicit prospective remembering when monitoring is eliminated before the target cue is presented. We gave participants either a single target word or initial-letter PM cue and presented the target cue at the end of a lexical decision block that contained many trials (500 trials). We imposed many lexical decision trials prior to presentation of the PM target because Loft et al. (2008, Experiment 3) found that when PM target cues were never presented during an ongoing lexical decision task (comprising 640 lexical decision trials), task interference was reduced in the last 100 trials relative to a condition in which the target cues were frequently presented during the ongoing task. Based on that finding, we expected that task interference would be significantly reduced (if not completely eliminated) following several hundred trials in which the cue did not occur. If word cues, but not initial-letter cues, trigger retrieval of the PM intention when monitoring is not (or minimally) engaged (i.e., after 500 ongoing trials without presentation of a PM target event), such a result would reinforce the conclusion of Experiment 3 that monitoring is necessary for nonfocal PM cues but that focal cues can spontaneously trigger PM retrieval.

One final feature of this experiment merits mention. Smith et al. (2007) argued that manipulating PM and control blocks within subjects will mask task interference if there are carryover effects from the PM block to the control block (but see Einstein & McDaniel, in press, for a counterargument). To ensure that the conclusions drawn from Experiment 3 could not be explained by this simple design factor, in the present experiment, following Smith et al., we included a between-subjects control group that never received the PM task.

Method

Participants and design. One hundred and twenty Washington University undergraduates participated for class credit or monetary compensation. To obtain power to detect task interference that was similar to that in the Smith et al. (2007) experiments, we used 40 participants in each PM target type (word cue, initial-letter cue, no-PM control) condition.

Procedure. In the same manner as in Experiment 3, participants first received the instructions to the lexical decision task followed by practice. Then they performed a control block consisting of 100 lexical decision trials. Following the control block, some participants received the PM task instructions. In the word condition, participants encoded a single target word (either *crossbar*, *blinding*, *void*, *wobbling*, *plagued*, *loophole*, *equator*, or *skunks*) and were instructed to press the *Q* key if they saw that particular word. In the initial-letter condition, the PM task was to press the *Q* key if one saw a word that began with a particular letter (i.e., participants encoded either the initial-letter *c*, *b*, *v*, *w*, *p*, *l*, *e*, or *s*). Participants in the initial-letter condition were instructed that the target initial letter would never occur on a nonword trial. All participants were reminded that their primary goal was to perform the lexical decision task as quickly and as accurately as possible and were asked to write down their task instructions on a sheet of paper before continuing (accuracy confirmed by experimenter). The control group did not receive the PM task instructions.

Participants next performed a set of computerized Raven's Progressive Matrices (Raven, Raven, & Court, 1998) trials for 5-min. After the delay, participants performed 511 lexical decision trials. The target item appeared only on trial 501 and was the same for the word and initial-letter condition (e.g., the target initial-letter *c* occurred in the word *crossbar*).

Materials. All filler items were generated from the Balota et al. (2007) English lexicon database. Filler items were 4–8 letters and 1–3 syllables in length and had a log-transformed Hyperspace Analogue to Language (HAL) frequency from 5.0 to 7.0. The eight target words were randomly chosen from the list of words generated from the database (see Appendix B). All items appeared in lowercase letters, and no filler words were forwardly associated with any of the target words (according to the Nelson et al., 1998, norms).

A reviewer raised the concern that our PM target words and filler words might have differed in frequency values (determined by HAL value). We list target and (averaged) filler word frequencies, as well as other word characteristics (number of letters and syllables) in Appendix B. When we compared all (eight) target words to all (250) filler words, a significant difference obtained in frequency, $t(256) = 2.70$. However, a close inspection of Appendix B reveals that the significant difference is driven by the word *void* (HAL value > 10), which was mistakenly used instead of the word *voids*, the latter which has a more moderate HAL value (5.60). When the word *void* was excluded, the target and filler words did not differ in frequency ($t < 1$), nor did they differ in number of letters, $t(255) = 1.89$, or number of syllables ($t < 1$). Importantly, none of the statistical tests in the present experiment was altered when we excluded the participants who received the *void* (*v* in the initial-letter condition) target (resulting in 35 participants being retained in each between-subjects condition).

Results

RT trimming was conducted in the same manner as in Experiment 3.

Control-block performance. We first confirmed that there were no accuracy or RT differences between the conditions on the initial control block (100 lexical decision trials). A 3 × 2 mixed ANOVA that included PM target type (word, initial letter, or control) as a between-subjects variable and lexical decision trial (words or nonwords) as a within-subjects variable produced a main effect of lexical decision trial for RT, $F(1, 117) = 47.03$, $MSE = 1{,}824.73$, $\eta_p^2 = .45$, and accuracy, $F(1, 117) = 25.32$, $MSE = .006$, $\eta_p^2 = .30$. These effects emerged because responding was slower (see Table 3) but more accurate on nonword trials ($M = 0.92$) than word trials ($M = 0.87$), thereby signaling a speed/accuracy tradeoff for lexical decision type. More importantly, there was no effect of PM target type condition on RT ($F < 1$) or accuracy ($F < 1$; $M_{control} = .90$, $M_{word} = .89$, and $M_{letter} = .90$), and no PM Target Type × Lexical Decision Trial interaction for RT, $F(2, 117) = 1.19$, $MSE = 1{,}824.73$, $\eta_p^2 = .02$ or accuracy ($F < 1$). Therefore, there were no between-subjects differences prior to the PM target type manipulation.

PM block performance across quintiles. Because the primary research question was whether participants are consistently monitoring throughout the task and whether the PM action is executed when task interference is absent, we next investigated task interference across trial sets of the experimental block. In the absence of discrete segmentations of a task block (e.g., Experiment 3's quartile sets were segmented by target events), we took an approach similar to that used by Loft et al. (2008) by examining 100-trial sets of the experimental block. Consistent with Loft et al., our primary interest was in determining whether task interference changed from the first 100 trials to the last 100 trials (i.e., the trials preceding the target event). For the dependent measure of proportion of correct lexical decision trials, we conducted a 3 × 2 × 2 mixed ANOVA that included PM target type (word, initial letter, or control) as a between-subjects variable and lexical decision trial (words or nonwords) and quintile (1 or 5) as within-subjects variables. There was a significant quintile main effect, $F(1, 117) = 29.20$, $MSE = .004$, $\eta_p^2 = .20$ (*M*s = .88 and .85 for Quintiles 1 and 5, respectively), and a significant lexical decision trial main effect, $F(1, 117) = 61.25$, $MSE = .004$, $\eta_p^2 = .34$ (*M*s = .82 and .91 for words and nonwords, respectively). There were no other

Table 3
Mean Trimmed Reaction Times (in Milliseconds) on the Lexical Decision Task as a Function of the PM Target Type, Lexical Decision Trial, and Quintile in Experiment 4

	No-PM control		Word PM cue		Initial-letter PM cue	
Quintile	Words	Nonwords	Words	Nonwords	Words	Nonwords
Control	648 (95)	690 (102)	627 (74)	673 (94)	657 (109)	683 (100)
1	650 (76)	661 (74)	649 (77)	684 (103)	708 (116)	734 (141)
2	663 (90)	670 (85)	664 (108)	669 (108)	720 (163)	708 (132)
3	659 (84)	671 (82)	647 (94)	665 (96)	697 (154)	704 (149)
4	652 (79)	645 (73)	653 (94)	655 (105)	698 (145)	683 (124)
5	662 (93)	683 (84)	664 (97)	688 (111)	690 (126)	708 (144)

Note. Standard deviations are in parentheses. PM = prospective memory.

significant effects (next largest $F[2, 117] = 2.12$, $\eta_p^2 = .02$, for the Quintile × Condition interaction). Importantly, there was not a PM target type main effect, $F(2, 117) = 1.08$, $MSE = .004$, $\eta_p^2 = .02$, as lexical decision accuracy was similar in the control (.87), word (.85), and initial-letter (.87) conditions.

We next examined mean trimmed RTs in a 3 × 2 × 2 mixed ANOVA that included PM target type (word, initial letter, or control) as a between-subjects variable and lexical decision trial (words or nonwords) and quintile (1 or 5) as within-subjects variables.[3] Responding was faster on word trials than nonword trials (see Table 3 for means), $F(2, 117) = 13.84$, $MSE = 1{,}490.29$, $\eta_p^2 = .19$. Most importantly, there was a significant PM Target Type × Quintile interaction, $F(1, 117) = 3.76$, $MSE = 1{,}490.29$, $\eta_p^2 = .03$ (next largest $F[2, 117] = 2.84$, $\eta_p^2 = .05$, for the PM target type main effect). Planned comparisons demonstrated slower responding in the initial-letter condition than in the control condition during Quintile 1, $F(1, 117) = 58.13$, $MSE = 1{,}490.29$, $\eta_p^2 = .42$, and Quintile 5, $F(1, 117) = 9.47$, $MSE = 1{,}490.29$, $\eta_p^2 = .11$, as well as slower responding during the initial-letter condition relative to the word condition during both Quintile 1, $F(1, 117) = 39.90$, $MSE = 1{,}490.29$, $\eta_p^2 = .25$, and Quintile 5, $F(1, 117) = 7.14$, $MSE = 1{,}490.29$, $\eta_p^2 = .06$. In contrast, there was no mean RT difference between the control and word conditions during either Quintile 1, $F(1, 117) = 1.71$, $MSE = 1{,}490.29$, $\eta_p^2 = .01$, or Quintile 5 ($F < 1$). This result replicates the overall task interference finding in Experiment 3 of greater overall task interference for initial-letter PM targets than word PM targets during a lexical decision task. This finding was predicted by the focal/nonfocal hypothesis but not the monitoring difficulty hypothesis.

Furthermore, despite substantial power to detect medium-sized (>.99), medium- to small-sized (.97), and small-sized (.73) effects, there was no difference in RTs from Quintile 1 to Quintile 5 in the control condition, $F(1, 117) = 3.89$, $MSE = 1{,}490.29$, $\eta_p^2 = .03$, or in the word condition, $F(1, 117) = 1.16$, $MSE = 1{,}490.29$, $\eta_p^2 = .01$. However, there was a significant speed-up from Quintile 1 to Quintile 5 in the initial-letter condition, $F(1, 117) = 6.61$, $MSE = 1{,}490.29$, $\eta_p^2 = .05$, thereby demonstrating a reduction in monitoring in this condition. Thus, as illustrated in Table 3, the PM Target Type × Quintile interaction obtained because there was less task interference in Quintile 5 than in Quintile 1 in the initial-letter condition and no significant task interference during either quintile in the word condition.

Because of the importance of the Quintile 5 analysis in determining whether task interference was present immediately preceding the target event, we conducted an additional analysis of covariance on Quintile 5 mean RTs in which PM target type (word, initial letter, or control) was a between-subjects factor, lexical decision type (words or nonwords) was a within-subjects factor, and word and nonword control block mean trimmed RTs were controlled. There were no significant effects (largest $F[2, 115] = 1.83$, $\eta_p^2 = .03$, for PM target type main effect) except for the planned comparison between the initial-letter condition and the control condition (calculated on adjusted Quintile 5 means), $F(1, 115) = 10.50$, $MSE = 2{,}124.60$, $\eta_p^2 = .08$. Both the initial-letter and word condition comparison, and the word and control condition comparison were not significant, $F(1, 115) = 3.13$, $MSE = 2{,}124.60$, $\eta_p^2 = .03$, and $F(1, 115) = 2.16$, $MSE = 2{,}124.60$, $\eta_p^2 = .02$, respectively. Thus, even when controlling for control block RTs, we found that task interference was not present in the word condition in the quintile before the target event.

PM task performance. The multiprocess and PAM theories predict different patterns of PM performance based on the Quintile 5 task interference results. According to the PAM theory, task interference is required for prospective remembering. Therefore, given that word and initial-letter cues are equivalent in monitoring

[3] For the sake of completeness, we also examined lexical decision performance during the middle 300 trials (Quintiles 2–4; see Table 3). We conducted a 3 × 3 × 2 mixed ANOVA on mean trimmed RTs in which PM target type (word, initial letter, control) was a between-subjects factor and quintile (2, 3, or 4) and lexical decision trial (words or nonwords) were within-subject factors. There was a quintile main effect, $F(2, 234) = 10.52$, $MSE = 1{,}060.04$, $\eta_p^2 = .08$, that was qualified by a Quintile × Lexical Decision Trial interaction, $F(2, 234) = 5.27$, $MSE = 1{,}060.04$, $\eta_p^2 = .04$ (see Table 3; next largest $F[2, 117] = 2.19$, $\eta_p^2 = .04$, for the PM target type main effect). The planned comparisons showed greater RTs in the initial-letter condition than in the control condition during Quintile 2, $F(1, 117) = 43.44$, $MSE = 1{,}060.04$, $\eta_p^2 = .27$; Quintile 3, $F(1, 117) = 23.54$, $MSE = 1{,}060.04$, $\eta_p^2 = .17$; and Quintile 4, $F(1, 117) = 33.07$, $MSE = 1{,}060.04$, $\eta_p^2 = .22$. Likewise, there was a significant difference between RTs in the initial-letter and word conditions during Quintile 2, $F(1, 117) = 43.45$, $MSE = 1{,}060.04$, $\eta_p^2 = .27$; Quintile 3, $F(1, 117) = 37.27$, $MSE = 1{,}060.04$, $\eta_p^2 = .24$; and Quintile 4, $F(1, 117) = 24.71$, $MSE = 1{,}060.04$, $\eta_p^2 = .17$. Control and word condition RTs were similar in Quintile 2 ($F < 1$); Quintile 3 ($F[1, 117] = 1.57$, $MSE = 1{,}060.04$, $\eta_p^2 = .01$) and Quintile 4 ($F < 1$).

difficulty (Experiments 2a and 2b), this theory anticipates that PM performance should be greater in the initial-letter condition than in the word condition because task interference obtained in the initial-letter but not in the word condition. Furthermore, because there was no evidence for task interference in the word condition proximal to the target event, the PAM theory would expect that PM performance in this condition should be at floor levels.

In contrast, the multiprocess theory predicts that in the absence of task interference, PM performance will be better in a focal (word) condition than a nonfocal (initial-letter) condition. Because task interference in the initial-letter condition was very low proximal to the target event (i.e., in Quintile 5) compared to the high levels of task interference obtained in Quintile 1 (which we assume reflected fairly consistent monitoring for most participants), only a few participants may have still been monitoring for the initial-letter cue by the target event. According to the multiprocess (and PAM) theory, nonfocal PM performance is supported by monitoring processes and not spontaneous retrieval processes; hence, PM performance should be lower in the initial-letter condition than the word condition.

PM task performance was computed as the proportion of participants who remembered to press the *Q* key on the target event or the following trial. The results were consistent with the multiprocess theory's prediction that PM performance would be better in the word condition ($M = 0.73$) than in the initial-letter condition ($M = 0.18$), $\chi^2(1) = 24.44$, $\phi = .55$. These results demonstrate that when monitoring (as indexed by significant task interference) is reduced or eliminated, PM performance will be much greater when a cue is focally processed than when the cue is not focally processed.

Given that overall task interference was minimal or absent but large differences in PM performance were observed for word and initial-letter cues, an important remaining question concerns whether task interference was present for those individuals who remembered to press the *Q* key ($n_{letter} = 7$ and $n_{word} = 29$). For these participants, we examined mean RTs on the five trials preceding the target event, excluding only incorrect responses (analogous to the proximal trial analysis used by Scullin et al., 2010). Responding was similar in the control ($M = 700$ ms) and word ($M = 707$ ms) conditions but appeared slower in the letter condition ($M = 863$ ms). We therefore compared mean RTs on these five trials (collapsing across word and nonword trials) by conducting a 3 × 2 mixed analysis of covariance in which PM target type (word, initial letter, or control) was a between-subjects factor, lexical decision type (words or nonwords) was a within-subjects variable, and control block mean trimmed RTs were controlled. There was a significant main effect, $F(2, 76) = 4.26$, $MSE = 32{,}509.57$, $\eta_p^2 = .10$; RTs did not differ between the word and control conditions ($t < 1$), but responding in the initial-letter condition was significantly slower than responding in both the word, $t(34) = 2.39$, and control $t(45) = 2.97$, conditions. Thus, monitoring immediately before a target event was critical for nonfocal, but not focal, PM (note that strikingly similar results obtain with a syllable nonfocal cue; see Scullin et al., 2010, Experiment 2).

General Discussion

Einstein et al. (2005) reasoned that, in the context of a category judgment task, a word cue but not a syllable cue would be focally processed and thereby support PM retrieval through a spontaneous process (rather than through monitoring; see McDaniel & Einstein, 2007b). Their results were consistent with this hypothesis in that task interference was greater in the syllable target condition than in the word target condition. However, if syllable targets are more difficult to monitor for than are word targets, then differences in task interference may have been caused by differences in cue difficulty. The results of the present Experiment 1 confirmed that a viable interpretation of Einstein et al.'s results could be that syllables are simply more difficult to monitor for than are words.

Accordingly, in Experiments 2a and 2b we attempted to identify cues that would be fairly equivalent in monitoring difficulty. This allowed us to more convincingly distinguish between the multiprocess and the monitoring interpretation of the effects of focal versus nonfocal cues on task interference and PM performance (Experiments 3 and 4). Empirical examinations of monitoring under the scenario in which identification of a PM target is made prior to responding to the ongoing task (Experiment 2a), as well as a scenario that allowed for individual differences in monitoring strategy (Experiment 2b), converged on the same result: An initial-letter cue and a word cue were highly similar in monitoring difficulty.[4]

In Experiment 3, we embedded an initial-letter cue and a word cue in an ongoing task (lexical decision) that directed attention toward focal processing of the word cue but not the initial-letter cue. Despite matching cues on monitoring difficulty (Experiments 2a and 2b), the results revealed a difference in overall task interference between the word and initial-letter conditions, with the focal word cues displaying significantly less task interference than the nonfocal initial-letter cues. Further, in some quartiles, for the focal word condition there was not even nominal evidence for any task interference. For instance, prior to the onset of the first PM target, in the word condition, the mean task interference was −4 ms. This overall pattern disfavors the cue-monitoring difficulty hypothesis (in which equivalent levels of task interference would have been expected), but is consistent with the a priori prediction of the multiprocess theory that the focality of the PM cue will determine the degree to which monitoring is required for PM retrieval. More specifically, the idea is that PM retrieval is dependent on monitoring (as indexed by task interference) for nonfocal cues (initial-letter in the present situation) but not focal cues (word cue). Further supporting the idea that monitoring is not necessary for focal cues, there was no statistical PM performance difference when significant task interference preceded the focal target event compared to when task interference was absent. This result is at odds with the PAM theory's claim that monitoring or other preparatory processes are always required for a focal cue to stimulate retrieval of a PM intention (Smith et al., 2007).

Experiment 4 reinforced and generalized the above findings by including more initial-letter/word target pairs (instead of only using *g* and *generous*) and by using the between-subjects (no-PM) control group design that Smith et al. (2007) have argued provides the clearest evidence for the presence of task interference. With a design assumed to be highly sensitive to the presence of task

[4] Even after gaining additional power to detect medium-sized effects (>.99) by collapsing across Experiments 2a and 2b, mean trimmed filler trial RTs did not differ statistically for word blocks ($M = 720$ ms) and initial-letter blocks ($M = 733$ ms; $F < 1$).

interference (lexical decision ongoing task coupled with a between-subjects control group; Smith et al., 2007), the results dovetailed with those of Experiment 3. As anticipated by the multiprocess theory, participants appeared to rely on qualitatively different retrieval processes in the focal and nonfocal conditions. Two critical results converge on this conclusion. First, for the (nonfocally processed) initial-letter cue, significant task interference was found, which suggests that resources were devoted to monitoring for the initial-letter cue. By contrast, for the (focally processed) word cue, no significant task interference was present (implying that monitoring was not engaged; if it had been engaged, then task interference should have been present as it was for the initial-letter cue). Most critically, for the word (focal) cue there was no significant task interference during the quintile that immediately preceded the focal target event, yet the PM action was executed nearly 75% of the time. Importantly, these results demonstrate that focal PM cues can trigger retrieval in the absence of task interference. This result further supports the idea that spontaneous retrieval processes can deliver a PM intention to consciousness.

Second, under Experiment 4's conditions that discouraged sustained monitoring and resulted in minimal task interference, initial-letter cue PM performance dropped substantially relative to that observed in Experiment 3 (when task interference preceding target events was robust). This pattern is clearly consistent with the idea that monitoring (indexed by task interference) is associated with PM performance for nonfocal cues. For the focal-cue group, however, as noted above, the pattern was very different. Here, monitoring was absent by the target event onset, but PM performance was still high. These findings provide compelling evidence that monitoring or other resource-demanding preparatory processes are not necessary for focal prospective remembering.

In addition to implicating spontaneous retrieval processes, the results of the quintile analysis in Experiment 4 also suggest that individuals may be more likely to rely on spontaneous retrieval than on monitoring for performing real-world intentions. Unlike laboratory settings, long retention intervals (on the order of hours or days instead of minutes or seconds) are commonly interleaved between formation and execution of real-world intentions. The results of Experiment 4 demonstrate that participants may monitor some following intention formation, but unreinforced monitoring is minimized or discontinued after several minutes (by Quintile 5). This is consistent with the premise that if one forms an intention in the morning that should be executed at night, one will not monitor all day for that intention (see, e.g., Scullin & McDaniel, in press). Importantly, when the PM cue is focally processed, participants may rely on spontaneous retrieval processes to achieve high levels of prospective remembering. The practical implication is therefore that one should set up cues that are likely to be focally processed near the time in which the intention should be executed.

In sum, the present research provides converging support for the multiprocess theory of PM retrieval. The multiprocess theory argues that in the absence of monitoring processes (i.e., task interference), individuals may spontaneously retrieve an intended action if the cue is focally processed. Consistent with the multiprocess theory's account, high focal PM performance was achieved under conditions in which task interference was absent proximal to target events. These results are consistent with the intuitive notion that in the real world, one need not engage monitoring every time a PM intention is formed. Instead, individuals may rely on environmental cues (Kvavilashvili & Fisher, 2007) to spontaneously remind them to execute their PM intentions.

References

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., . . . & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459.

Breneiser, J. E., & McDaniel, M. A. (2006). Discrepancy processes in prospective memory retrieval. *Psychonomic Bulletin & Review, 13,* 837–841.

Cohen, A.-L., Jaudas, A., & Gollwitzer, P. M. (2008). Number of cues influence the cost of remembering to remember. *Memory & Cognition, 36,* 149–156.

Einstein, G. O., & McDaniel, M. A. (in press). 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.*

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science, 14,* 286–290.

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.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3 [Computer software]. Retrieved from http://www.psycho.uni-duesseldorf.de/aap/projects/gpower

Kvavilashvili, L., & Fisher, L. (2007). Is time-based prospective remembering mediated by self-initiated rehearsals? Effects of incidental cues, ongoing activity, age, and motivation. *Journal of Experimental Psychology: General, 136,* 112–132.

Loft, S., Kearney, R., & Remington, R. (2008). Is task interference in event-based prospective memory dependent on cue presentation? *Memory & Cognition, 36,* 139–148.

Loft, S., & Yeo, G. (2007). An investigation into the resource requirements of event-based prospective memory. *Memory & Cognition, 35,* 263–274.

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

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology, 14,* S127–S144.

McDaniel, M. A., & Einstein, G. O. (2007a). *Prospective memory: An overview and synthesis of an emerging field.* Thousand Oaks, CA: Sage.

McDaniel, M. A., & Einstein, G. O. (2007b). Spontaneous retrieval in prospective memory. In J. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roedgier III* (pp. 227–242). Hove, England: Psychology Press.

McDaniel, M. A., Einstein, G. O., & Rendell, P. G. (2008). The puzzle of inconsistent age-related declines in prospective memory: A multiprocess explanation. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 141–160). Mahwah, NJ: Erlbaum.

McDaniel, M. A., Guynn, M. J., Einstein, G. O., & Breneiser, J. E. (2004). Cue focused and automatic-associative processes in prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 605–614.

McDaniel, M. A., Robinson-Riegler, B., & Einstein, G. O. (1998). Prospective remembering: Perceptually driven or conceptually driven processes? *Memory & Cognition, 26,* 121–134.

McDaniel, M. A., & Scullin, M. K. (2010). Implementation intention encoding does not automatize prospective memory responding. *Memory & Cognition, 38,* 221–232.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University* *of South Florida word association, rhyme, and word fragment norms*. Retrieved from http://w3.usf.edu/FreeAssociation/

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Oxford, England: Oxford Psychologists Press.

Scullin, M. K. (2009). *Sleeping to remember: Spontaneous retrieval of prospective memories across sleep and wake delays.* Unpublished master's thesis, Washington University, St. Louis.

Scullin, M. K., Einstein, G. O., & McDaniel, M. A. (2009). Evidence for spontaneous retrieval of suspended but not finished prospective memories. *Memory & Cognition, 37,* 425–433.

Scullin, M. K., & McDaniel, M. A. (in press). Remembering to execute a goal: Sleep on it! *Psychological Science.*

Scullin, M. K., McDaniel, M. A., & Einstein, G. O. (2010). Control of cost in prospective memory: Evidence for spontaneous retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 190–203.

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.

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

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.

Appendix A

Target Items From Experiments 1–2b (Across Counterbalanced Lists)

Experiment 1		Experiment 2a and 2b
Word	Syllable	Initial letter
	List 1	
Verily*	Ver*	V*
Leisurely*	Sur*	L*
Unaware*	Ware*	U*
Frantically*	Fran*	F*
Eligible*	Gib*	E*
Intrepid	Tre	I
Yonder	Der	Y
Navigate	Gate	N
Turnpike	Turn	T
Wavelengths	Wave	W
	List 2	
Generous*	Gen*	G*
Amenable*	Men*	A*
Delivering*	Del*	D*
Pacifier*	Cif*	P*
Commanding*	Ding*	C*
Rambling	Ram	R
Healer	Ler	H
Operas	Per	O
Muster	Mus	M
Stillness	Ness	S

Note. An asterisk (*) signals a critical target item.

Appendix B

Prospective Memory Target-Word and Filler-Word Characteristics in Experiment 4

Prospective memory target word	No. of letters	No. of syllables	Frequency[a]
Individual target words			
Blinding	8	2	6.70
Crossbar	8	2	5.04
Equator	7	3	6.78
Loophole	8	2	6.42
Plagued	7	1	6.80
Skunks	6	1	6.19
Void	4	1	10.04
Wobbling	8	3	5.11
Averaged values			
Target words (including *void*)	7.00 (1.41)	1.88 (0.83)	6.63 (1.55)
Target words (excluding *void*)	7.43 (0.79)	2.00 (0.82)	6.15 (0.76)
Filler words	6.60 (1.15)	1.99 (0.63)	6.04 (0.57)

Note. Averaged values are given for target words (including vs. excluding *void*) and filler words. Standard deviations are in parentheses.
[a] Log-transformed Hyperspace Analogue to Language value.

Received December 19, 2008
Revision received January 7, 2010
Accepted January 11, 2010 ■