Journal of Experimental Psychology:
Learning, Memory, and Cognition
2004, Vol. 30, No. 4, 756–777

0278-7393/04/$12.00 DOI: 10.1037/0278-7393.30.4.756

A Multinomial Model of Event-Based Prospective Memory

Rebekah E. Smith and Ute J. Bayen
The University of North Carolina at Chapel Hill

Prospective memory is remembering to perform an action in the future. The authors introduce the 1st formal model of event-based prospective memory, namely, a multinomial model that includes 2 separate parameters related to prospective memory processes. The 1st measures preparatory attentional processes, and the 2nd measures retrospective memory processes. The model was validated in 4 experiments. Manipulations of instructions to place importance on either the prospective memory task or the background task (Experiments 1 and 2) and manipulations of distinctiveness of prospective memory targets (Experiment 2) had expected effects on model parameters, as did a manipulation of the difficulty of prospective memory target encoding (Experiments 3 and 4). An alternative model was also evaluated.

Remembering to perform an action in the future is referred to as prospective memory. For example, one may have to take medication at 10 p.m. or give one's colleague a message when one sees him or her. The former is an example of a time-based task. A time-based task refers to remembering to perform an action at a specific time or after a certain amount of time has elapsed. The latter example is an event-based task, one that must be performed when a certain event occurs. Both time-based and event-based prospective memory tasks constitute a crucial form of memory use in our daily lives, but despite the importance of this form of memory, research on prospective memory was minimal until recently (Einstein & McDaniel, 1996). The recent increase in interest in this area of research has resulted in the development of a number of different explanatory frameworks (e.g., Einstein & McDaniel, 1996; Ellis, 1996; Goschke & Kuhl, 1996; Guynn, McDaniel, & Einstein, 2001; McDaniel & Einstein, 2000; Smith, 2003). Our goal in the work presented here was to develop and evaluate a formal mathematical model for the investigation of event-based prospective memory.

The extant explanations of prospective memory rely heavily on latent cognitive processes. Unfortunately, the traditional methods of statistical analysis that have been used in prospective memory research cannot adequately address questions related to latent cognitive processes. Multinomial process tree (MPT) models are substantively motivated statistical models that provide a means for measuring latent cognitive processes by estimating model parameters that represent these processes from observable data (Batchelder & Riefer, 1999). The structure of an MPT model is determined by theoretically based assumptions about the relationships among various underlying latent cognitive processes. The modeler must specify exactly how the processes interact with one another. In the experiments presented in this article, we extend the use of MPT models to the area of prospective memory research and provide the first application of mathematical models to this area.

To illustrate how MPT models can contribute to this area of research, consider the distinction made previously by Einstein and McDaniel (1990, 1996) between the prospective component and the retrospective component of prospective memory performance. According to these authors, the retrospective component is the what and the when of a prospective memory task, whereas the prospective component is the coming to mind of the prospective memory response at the appropriate time (Einstein & McDaniel, 1996). Remembering that you have to do something is the prospective component, whereas remembering what you have to do and when you have to do it is considered the retrospective component. Researchers have disagreed about the exact nature of the prospective and retrospective components of event-based prospective memory. Unfortunately, if a variable affects prospective memory, we cannot determine how the variable affects each of the two different components using traditional accuracy measures.[1] This places limitations on our ability to investigate the nature of the two components and how they contribute to prospective memory performance. The goal is to use MPT models to disentangle the respective contributions of prospective and retrospective processes to task performance and thus measure how each of the two components is affected by experimental manipulations. The model is based on the preparatory attentional processes and memory processes (PAM) theory of prospective memory (Smith, 2003), which proposes that the prospective component involves processes that draw on our limited resources, and, thus, that these processes are

Rebekah E. Smith and Ute J. Bayen, Department of Psychology, The University of North Carolina at Chapel Hill.

The research reported in this article was supported by Individual National Research Service Award F32 AG20021 and by Grant R01 AG17456 from the National Institute on Aging. We thank Jessica Scorso for her assistance with data collection and Bill Batchelder, Mark McDaniel, David Riefer, and Julia Spaniol for helpful comments on drafts of this article.

Correspondence concerning this article should be addressed to Rebekah E. Smith, Department of Psychology, CB 3270, Davie Hall, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599. E-mail: rebekah@unc.edu

[1] Although prospective memory accuracy measures cannot distinguish between the effects of manipulations on the retrospective and prospective components, researchers have attempted to distinguish between the two by requiring different responses for the two components (e.g., A.-L. Cohen et al., 2001, 2003). Although this is a clever and useful solution for distinguishing between the two components, the additional response requirement creates a secondary intention, potentially altering the nature of the prospective memory task. With MPT models, we can measure these two components without adding an additional prospective memory response.

not automatic (Posner & Snyder, 1975). A brief description of the typical laboratory paradigm used to study event-based prospective memory is needed before we describe the evidence that indicates that nonautomatic processes are involved in prospective memory.

In humans' everyday lives, prospective memory tasks are accomplished while they are engaged in other activities. For instance, you decide in the morning that you need to stop at the store on the way home from work to buy milk. You probably do not sit in your office all day just waiting for the opportunity to carry out your intention; rather, you are engaged in other activities during this delay interval. Even during the performance interval, the time during which the task could be accomplished when the target event (seeing the store) occurs, you must devote resources to driving, attending to traffic, perhaps listening to the radio, and so on. You must stop when the target event occurs, rather than continuing to drive home as you would on a day when you do not have to perform this prospective memory task. In order to mimic the fact that in the real world, prospective memory tasks occur in the midst of other activities and involve interrupting those activities, laboratory prospective memory tasks are embedded in an ongoing activity. For instance, participants may need to remember to press a special key on a keyboard if a certain word occurs during a word-rating task (Einstein, Smith, McDaniel, & Shaw, 1997).

Using a similar paradigm, Smith (2003) showed that participants who have to perform an embedded event-based prospective memory task while also performing a lexical decision task take between 200 and 300 ms longer to respond to the lexical decision task than participants who performed the lexical decision task alone. Importantly, this difference in response time was found on nonprospective memory target trials. The cost to the ongoing task indicates that some sort of nonautomatic processes were being engaged to prepare for the prospective memory task before the target event occurred. Similar findings have been reported in a number of different studies (Burgess, Quayle, & Frith, 2001; Marsh, Hicks, Cook, Hansen, & Pallos, 2003), including cases in which a single salient target event was used (Smith & Hunt, 2004).

According to PAM theory, capacity-consuming preparatory processes must be engaged for successful event-based prospective memory. These preparatory attentional processes may include nonautomatic monitoring of the environment for the prospective memory-target events. Thus, because nonautomatic preparatory processes must be engaged during the performance interval (i.e., the interval in which the intended action can be performed), the prospective component of the task, that is, the retrieval of an intention, is not automatic. The preparatory processes occur prior to the occurrence of the target event. Therefore, when a prospective memory task is embedded in an ongoing task, this will reduce the resources available for the ongoing task, even when the target event is not present. The preparatory attentional processes are functionally related to prospective memory performance. That is, better prospective memory performance should be accompanied by increased monitoring, which will come at a greater cost to the ongoing activity, when comparing performance in the same task situation.

The preparatory processes are likely to be complex and dependent on the particular task demands. For instance, beyond the nonautomatic monitoring, preparatory processes may include rehearsal of the prospective memory-target events. Thus, variations in the number and nature of prospective memory targets could also contribute to variations in the resources needed for engaging in sufficient preparatory processing for successful prospective memory performance. Furthermore, the ability to engage in preparatory processing will be dependent on the extent of resources available to an individual and the extent to which the ongoing task is resource demanding. The likelihood of engaging in preparatory processing can also be influenced by the importance of the task and by nontargets that are highly similar to the targets, thereby serving as potential cues for the prospective memory task. The effects of task importance and target–nontarget similarity are considered in this article.

In addition to the preparatory processes, according to the PAM theory, retrospective memory processes are also involved in prospective memory performance. Retrospective memory processes are needed for discrimination between prospective memory target and nontarget events, as well as for recollection of the intended action. These memory processes are likely to absorb resources when the target is present. Marsh, Hicks, and Watson (2002) found evidence that nonautomatic processes are engaged when the target event is present. The likelihood of correctly recognizing the target events can be affected by the number of target events, how well the targets are encoded, and the relationship between target and nontarget events in the ongoing task. We investigated the latter two effects in this study.

We believe that these two components, preparatory attentional processes and retrospective memory processes, interact to determine whether a prospective memory task will be accomplished. The PAM theory is distinguished from extant explanations by the proposal that resource-demanding preparatory attentional processes are always required for successful prospective memory performance (Smith, 2003). For instance, according to the multiprocess framework (McDaniel & Einstein, 2000), event-based prospective memory tasks involve varying degrees of strategic processing. According to this view, some event-based prospective memory tasks may require extensive strategic processing, whereas others can be accomplished in a more automatic fashion (Kliegel, Martin, McDaniel, & Einstein, 2001; McDaniel & Einstein, 2000). The PAM theory underlies the multinomial model that we now present.

The Multinomial Model

MPT models are a family of formal models for categorical data. In these models, it is assumed that there are discrete cognitive states that participants attain with certain probabilities during task performance. These probabilities are represented as model parameters that can be estimated from observed raw data via maximum-likelihood parameter estimation. The fit of the resulting model to the empirical data can be evaluated via goodness-of-fit tests (see Hu & Batchelder, 1994, or Riefer & Batchelder, 1988, for technical details). Multinomial models have been enjoying increasing popularity in cognitive psychology over the last 20 years (for reviews, see Batchelder & Riefer, 1999; Erdfelder, in press) with applications for a diversity of cognitive tasks.

When developing a new model, it is important not only to evaluate its goodness of fit to empirical data but also to perform experimental validation (Batchelder & Riefer, 1999). In experimental validation studies, basic independent variables are manipulated and shown to have predicted selective effects on model parameters. For examples of this approach, see Bayen, Murnane, and Erdfelder (1996); Buchner, Erdfelder, and Vaterrodt-Plünnecke (1995); Erdfelder and Buchner (1998); and Yu and Bellezza (2000). After model validation has been successfully performed, the model can be applied to test theories (see, e.g., Bayen, Nakamura, Dupuis, & Yang, 2000; Spaniol & Bayen, 2002) or to investigate cognitive processes in different populations (e.g., Bayen & Murnane, 1996).

In accord with the PAM theory described above, our multinomial model includes separate and independent parameters for preparatory attentional processes and for retrospective memory processes. We have assessed the validity of the model in four experiments involving experimental manipulations that should differentially affect these different model parameters.

The current model is designed for ongoing tasks with two response alternatives and two trial types. In the experiments presented here, for the purposes of validating the model, we used a color-matching task as the ongoing activity. In the color-matching task, participants saw four rectangles, each displayed in a different color, presented sequentially. After the fourth rectangle, participants saw a word printed in a color. The color of the word was either the same as one of the four color rectangles (a match trial) or a different color (a nonmatch trial). The participants' task was to decide whether the color of the word matched one of the four colors that had just been shown, and to make their responses by pressing either the *Y* key (for *yes*) or the *N* key (for *no*). For the prospective memory task in the current experiments, participants tried to remember to press the tilde key when certain words occurred in the color-matching task. Thus, the color-matching task with an embedded event-based prospective memory task had four trial types: prospective memory targets in which the color matched one of the previous set of four colors (target, match trials), prospective memory targets in which the color did not match one in the previous set (target, nonmatch trials), nonprospective memory target trials in which the color matched (nontarget, match trials), and nonprospective memory target trials in which the color did not match (nontarget, nonmatch trials). In Experiments 1–3, each trial had three possible responses: *Y*, *N*, or the prospective memory response (pressing the tilde key).

Our MPT model for three response classifications is illustrated in Figure 1. When a prospective memory-target word is presented on a match trial (the top tree in Figure 1), C_1 is the probability that the color will be detected as a match with one of the colors in the previous set. The probability that a participant will engage in preparatory attentional processes is represented by *P*, regardless of whether the color of the target is a match. Recognition of the prospective memory target, probability M_1, results in a prospective memory key response. If the participant is engaging in preparatory processes, but fails to recognize the target (probability $1 - M_1$), he or she may still guess that the word is a target with probability *g*. Alternatively, the participant may not guess that the word is a target ($1 - g$) and say that the color of the word is a match and respond by pressing the *Y* key. If the participant is not engaging in preparatory processes ($1 - P$), the word is identified as a match, resulting in a *Y* keypress. The bottom half of the tree represents the case in which the participant does not recognize the color as a match ($1 - C_1$). In this case, a prospective memory keypress will occur under the same circumstances as described when the color match is detected. The participant must be engaging in preparatory

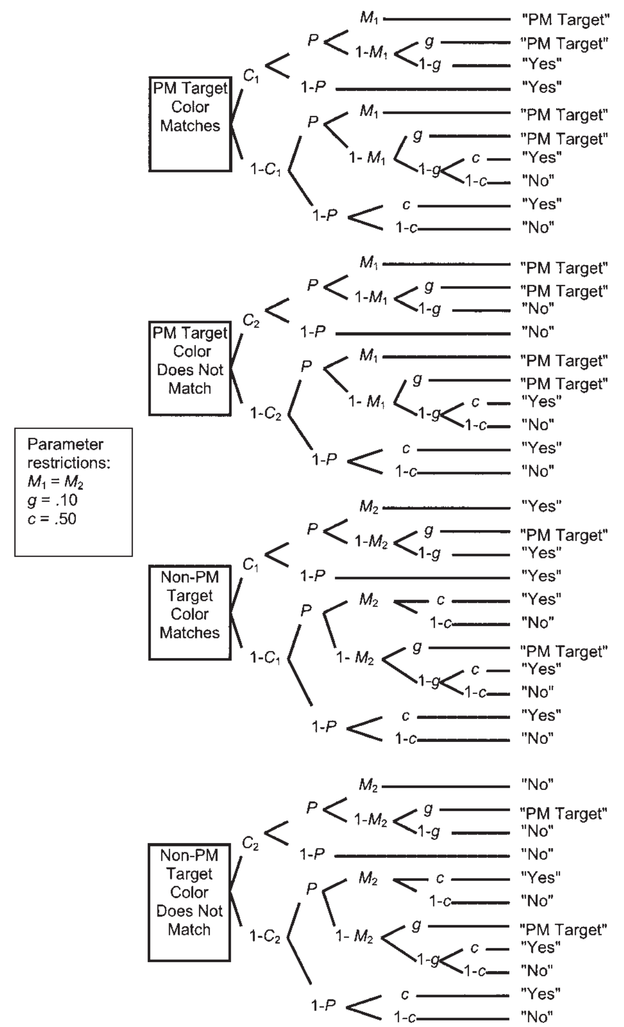

Figure 1. Multinomial model of event-based prospective memory based on the preparatory attentional processes and memory processes theory. PM = prospective memory; C_1 = probability of detecting a color match; C_2 = probability of detecting that a color does not match; P = preparatory attentional processes; M_1 = probability of detecting a PM target; M_2 = probability of detecting that a word is not a PM target; g = probability of guessing that a word is a target; c = probability of guessing that a color matches.

processes, *P*, and then either recognize the item as a target (M_1) or, if the item is not recognized (with probability $1 - M_1$), guess that it is a target (*g*) in order for a prospective memory keypress to occur. If the target is not recognized ($1 - M_1$), and the participant does not guess that the word is a target ($1 - g$), then the participant may guess that the color matches with probability *c*, resulting in a *Y* keypress, or the participant may guess that the color does not match, ($1 - c$), resulting in an *N* keypress. If the participant is not engaging in preparatory processes ($1 - P$), then he or she could either guess that the color matches, *c*, and press the *Y* key, or not guess that the color matches, $1 - c$, and press the *N* key. The second tree, representing target words on nonmatch trials, is the same as the first tree, except for C_2, the probability of recognizing the color as not matching one from the previous set.

The third tree and fourth tree illustrate nontarget match trials and nontarget nonmatch trials, respectively. The PAM theory proposes that preparatory processes must occur and are strategically determined by the participant. Until those preparatory processes occur, the retrospective memory component will not come into play. That is, prior to the occurrence of preparatory processes, the participant cannot distinguish between target and nontarget trials. Thus, the preparatory processes should occur for both target and nontarget trials and should be equal for both types of trials. Therefore, as in the first two trees, P is the probability that preparatory processes will be engaged, regardless of the color-match status. The probability of recognizing that a nontarget word is not a target is designated by M_2. Note that the model makes a two-high-threshold (2HT) assumption regarding item recognition. With probability M_1, target items cross a threshold for target recognition. A second threshold is for nontarget items. With probability M_2, nontargets cross this threshold and are then known to be nontargets. 2HT models of item recognition are adequate in predicting empirical receiver operating characteristics in other paradigms (Macmillan & Creelman, 1990; Snodgrass & Corwin, 1988; Swets, 1986). A multinomial model of source monitoring that included a 2HT assumption for item recognition was presented by Bayen et al. (1996) and passed empirical validity tests. It predicted data better than alternative models with different threshold assumptions. Note that our model also includes a 2HT assumption regarding color recognition: C_1 indicates the probability that a matched color crosses a recognition threshold, whereas C_2 indicates the probability that a nonmatched color crosses a threshold and is correctly recognized as a nonmatch.[2]

In the third and fourth trees of the model, a prospective memory keypress will occur only if the participant is engaging in preparatory processes, P, but fails to recognize that a word is not a target $(1 - M_2)$ and guesses, with probability g, that it is a target. When a nontarget word is recognized as a nontarget word, M_2, participants will respond Y if they detect that a color match trial is a match, with probability C_1, or if they fail to detect that a nonmatch trial color does not match $(1 - C_2)$, and then guess, with probability c, that the trial is a match. Participants will respond N if they detect that a match trial is not a match (probability C_2), or if they fail to detect that a match trial color does match $(1 - C_1)$, and then guess, c, that the trial is a match. As proposed by the PAM theory, prospective memory responses cannot occur in the absence of preparatory attentional processes.

The probability of a certain response category is the sum of the branches in the processing tree that lead up to that response. Model parameters are estimated from the resulting model equations. The seven-parameter version of the model that is shown in Figure 1 is not identifiable (see Appendix A). Therefore, we imposed theoretically motivated constraints on model parameters. Such constraints may be either equality constraints setting two (or more) parameters equal to each other (see, e.g., Batchelder & Riefer, 1990; Bayen et al., 1996) or constraints that set parameters to certain predetermined values (Erdfelder & Buchner, 1998). Constraints on our model parameters lead to an identifiable and testable four-parameter submodel (see Appendix A for the proof of identifiability). The constraints are $M_1 = M_2$; $c = .50$; and $g = .10$, resulting in a model with the four free parameters P, M, C_1, and C_2. The theoretical motivations for our parameter constraints are explained in turn.

The memory component M reflects a process of discriminating between target and nontarget items. We assume that the ability to detect targets is equal to the ability to detect that nontargets are not targets, in other words, $M_1 = M_2$. This is a standard assumption for the 2HT model of simple old–new item recognition. Researchers have successfully made the same assumption in studies using a multinomial model of source monitoring that also includes a 2HT item-recognition component (e.g., Bayen et al., 1996, 2000; Spaniol & Bayen, 2002; Yu & Bellezza, 2000).

We followed Batchelder and Riefer's (1999) recommendation to place constraints on ancillary parameters in order to leave crucial parameters free to vary. We set the ancillary response-bias parameters g and c equal to particular values. Participants sometimes match their response biases to the perceived ratio of different item types at test (e.g., Buchner et al., 1995; Ratcliff, Sheu, & Gronlund, 1992; Van Zandt, 2000). This is known as probability matching. Klauer and Wegener (1998) set guessing parameters in a multinomial model of source monitoring to specific values based on the assumption of probability matching. In our current experiments, we set the probability of guessing that a word is a target at the ratio of target to nontarget trials, $g = .10$, and set the probability of guessing that a color-match trial matches, c, to .50, because in our experiments, half of the trials matched, and half did not match. We assume that parameters g and c do not depend on prior memory and decision processes. For example, parameter g is assumed to be the same, regardless of whether on any given trial a color match or nonmatch is detected. This assumption appears reasonable, because probability matching should be independent of the detection of the color status of an item and results in $g = .10$ for both match and nonmatch trials.[3] Note that our parameter constraints and the corresponding assumptions are not arbitrary. These assumptions are theoretically motivated, and the goodness of fit of models that include these assumptions to empirical data can be tested.

[2] The fits of 1HT models were evaluated for the data from each group of participants in all four experiments by individually setting M_2 and C_2 equal to zero. The values of the test statistic $G^2(4)$ ranged from 41.26 to 894.34, all larger than the critical value of 9.49, indicating that the 1HT models do not fit these data. The 1HT versions of the four-response model in Experiment 4 were evaluated in the same fashion. The values of the test statistic $G^2(8)$ for the four-response model ranged from 44.15 to 618.21, all larger than the critical value of 15.51, indicating that the 1HT models do not fit these data either.

[3] We evaluated the sensitivity of the parameters of primary interest, P and M, to changes in the values of g and c in each condition for all experiments. When g is set to values between .06 and .33, the estimates of P do not differ significantly from the estimates reported in Experiments 1–3. Parameter M does not differ significantly when g is set to values between .06 and .17. In the three-response model in Experiment 4, values of P are stable across values of g between .06 and .23, whereas M is stable when g ranges from .07 to .17. For the four-response model, values of P are unchanged when g varies between .07 and .18, and M is unchanged when g varies between .07 and .15. In all cases of the three-response model, both P and M are stable over values of c ranging from .07 to .85. In the four-response model, P is stable when c ranges from .09 to .85, whereas M is stable for values of c between .38 and .85. Thus, the parameter estimates related to the prospective memory task are fairly robust against violations of the probability-matching assumptions.

Our goal in the current experiments was to validate the model, which involved manipulation of variables that should influence crucial model parameters in predictable and separable ways. We manipulated variables that had independent effects on the preparatory-processes parameter, *P*, and the memory-processes parameter, *M*. In Experiment 1, parameter *P* was influenced by a manipulation of task importance, whereas parameter *M* remained constant between groups. In Experiment 2, experimental manipulations affected these two parameters in predictable and opposite directions, and we also replicated the effect shown in Experiment 1. In Experiments 3 and 4, a manipulation of target encoding time affected the memory parameter without affecting the preparatory attentional processes parameter.

Experiment 1

In Experiment 1, participants completed two blocks of the color-matching task described in the introduction. A control group performed only the color-matching task in both blocks. Two groups of participants performed one block of the color-matching task alone and one block with the embedded event-based prospective memory task. In the prospective memory task, participants were to remember to press a special key if they saw one of six prospective memory target words during the color-matching task. One group was told that the prospective memory task was more important than the ongoing color-matching task. The other group was told that the color-matching task was more important than the prospective memory task. Prospective memory tasks that are said to be important are more likely to be performed (Andrzejewski, Moore, Corvette, & Herrmann, 1991; Krishnan & Shapiro, 1999; Kvavilashvili, 1987; but see also Kliegal et al., 2001). On the basis of these earlier findings and the PAM theory, we predicted that the Prospective Memory Importance (PMI) group would have better prospective memory performance; that is, they would be more likely to remember to perform the intended action when a target event occurred than would the Color-Matching Importance (CMI) group. Furthermore, the prospective memory importance should also show a greater cost on the ongoing activity (as measured by reaction times on nonprospective memory target trials), and the extent of this cost should be correlated with prospective memory performance. Similarly, Smith (2003) found that participants who were more likely to correctly perform an embedded prospective memory task were slower to make their responses on nonprospective memory target trials of the ongoing task than were participants who were less likely to remember to perform the prospective memory task. Some participants may have considered the prospective memory task to be more important than did other participants. The former participants may have engaged in more preparatory processing, thus improving their performance, but at a greater cost to the ongoing task. Because this increase in cost to ongoing-task performance occurred on trials that preceded target events, this increase was, arguably, due to preparatory processes. Thus, the manipulation of task importance should affect the multinomial model parameter *P*, which measures the likelihood of engaging in preparatory processes. All participants in Experiment 1 learned the targets to criterion prior to starting the prospective memory task in order to establish consistent target encoding across groups. Given that we controlled target encoding, we predicted that the parameter *M*, which measures the retrospective memory processes that contribute to prospective memory performance, would not be influenced by the manipulation. Therefore, we predicted that Experiment 1 would show that the manipulation of task importance affects the *P* parameter but leaves the *M* parameter unaffected.

Method

Participants

Participants in all four experiments were native English speakers with normal color vision who completed this study in exchange for partial course credit in introductory psychology classes. In all experiments, the sessions lasted approximately 30 min, and each session included between 1 and 5 participants.

Design

In this experiment, 82 participants completed two blocks of color-matching trials. Participants were randomly assigned to one of three instruction groups. Thirty-two participants received the PMI instructions, which emphasized the importance of the prospective memory task, and 32 received the CMI instructions, which emphasized the importance of the color-matching task. Eighteen participants were assigned to the control group and did not perform the prospective memory task. The control group was included to examine the effects of fatigue on response times to the color-matching task in the absence of an embedded prospective memory task. Trial type (color match vs. nonmatch) was manipulated within subjects.

Materials

A total of 124 medium-frequency words were randomly selected from the Kučera and Francis (1967) norms. From this larger set, two sets of 6 prospective memory target words were selected. The two sets were matched for frequency and word length. The remaining 112 words, which served as filler items in the ongoing task, were randomly assigned to each of two filler lists. Thus, there were two possible 6-item target word lists and two 56-item filler lists, one for each block of the ongoing task, and two possible orders for each target list and each filler list. In Order A, List 1 occurred in the first block and List 2 in the second block. In Order B, List 2 occurred first, followed by List 1. The resulting two possible counterbalancing combinations occurred equally often in the two prospective memory conditions. In the control group, one combination occurred 8 times, and the other occurred 10 times. Each target list served as the prospective memory target list equally often in each prospective memory condition.

The six target words occurred on Trials 10, 20, 30, 40, 50, and 60. We used equal intervals between targets to maximize the distance between each target. It is not uncommon to use evenly spaced targets in prospective memory research (e.g., Marsh, Hancock, & Hicks, 2002; Maylor, Smith, Sala, & Logie, 2002; McGann, Ellis, & Milne, 2002, 2003; Reese & Cherry, 2002; Smith, 2003). The filler words were randomly assigned to the other 56 trials for a total of 62 trials in each block. The targets and filler words were indistinguishable from the participants' perspective in the first block for all participants and in the second block for participants in the control condition.

Five colors were selected for the color-matching task: blue, green, red, white, and yellow. Color rectangles of 83 $\times$ 60 pixels (approximately 1.5 $\times$ 1.3 in. [3.81 $\times$ 3.30 cm]) were displayed in the center of a black screen. Words were displayed in the center of the screen in an 18-point font in one of the five colors on a black screen.

Procedure

Presentation of instructions and materials as well as response collection for this and all following experiments were computer directed. Participants read the color-matching task instructions, which emphasized both speed and accuracy. Participants then completed six practice trials, three matching and three nonmatching. Following the practice trials, participants had the opportunity to ask questions, and the experimenter reviewed the instructions.

Each color-matching trial included four color rectangles displayed individually in the center of a black screen for 500 ms each. Every color rectangle was followed by a completely blank black screen for 250 ms. Each of the four rectangles was a different color. Following the fourth color rectangle and blank screen, a lowercase word was displayed in the center of a black screen. On half of the trials (match trials), the color of the word was one of the four colors shown in the preceding set (correct response, *Y*). On the other half of the trials (nonmatch trials), the word was shown in the color that had not been shown in the set (correct response, *N*). After the participant responded with either the *Y* or the *N* key, the word was removed from the screen, and a blank screen was displayed for 1 s, followed by the first color rectangle of the next trial. There were 62 color-matching trials in each block. The order of match and nonmatch trials was determined randomly, as was the selection of colors for each trial.

At the end of the first block of color-matching trials, participants read that they had completed the first part of the color-matching task and that they would be doing other tasks for a while before returning to the second part of the color-matching task. Participants in the control group received no other instructions and went on to the delay activities described below. Participants in the prospective memory groups studied the six target words to a criterion of perfect recall. That is, the targets were displayed simultaneously on the computer screen until participants indicated that they had finished studying the words. Then the words were removed, and participants tried to recall them by writing them on a small slip of paper in any order. If a participant could recall all words, the experimenter started the next block of color-matching trials. If the participant could not recall all of the target words, they were displayed a second time. None of the participants took more than two attempts to recall all six words. The slip of paper on which the participants wrote the target words was removed before the participant continued.

Participants in the prospective memory groups then read the appropriate prospective memory instructions, which emphasized the importance of either the color-matching task (CMI group) or the prospective memory task (PMI group). The prospective memory action was to press the tilde key when one of the words appeared in the color-matching task. Instructions did not specify whether the tilde key alone should be pressed or whether (in response to a target word) both the tilde key and the *Y* or *N* key should be pressed. If participants asked, they were told that they could press the tilde key only or they could press the tilde key after having pressed the *Y* or *N* key. All responses were recorded, regardless of when the response was made. No participants in Experiment 1 pressed the tilde key after pressing the *Y* or *N* key. Participants were given the opportunity to ask questions, and the experimenter reviewed the instructions, emphasizing either the prospective memory task or the color-matching task.

A delay between the prospective memory instructions and the start of the performance interval can reduce prospective memory performance (Brandimonte & Passolunghi, 1994). The filled delay was included to avoid ceiling effects in prospective memory performance. During the delay interval, which lasted approximately 9 min, participants completed a computerized self-paced multiple-choice vocabulary test. Following the vocabulary test, participants performed a computerized letter-comparison task (Salthouse & Babcock, 1991).

After the letter-comparison task, participants began the second block of color-matching trials. They were not reminded about the prospective memory task. Three of the prospective memory targets occurred on match trials, and three targets occurred on nonmatch trials. Order of target occurrence and trial type (match vs. nonmatch) was random. Following the second block of color-matching trials, participants in the two prospective memory conditions recalled which key they were supposed to press in response to a target word and attempted to recall the six target words. Three participants who failed to correctly recall the key they were supposed to press were replaced. Participants also reported which task they believed to be most important. Four participants in the PMI condition reported that they thought that the color-matching task was most important, and 1 participant in the CMI condition reported that he or she thought the prospective memory task was most important. These individuals were replaced.

Results and Discussion

Prospective Memory Accuracy and Target Recall

We set an alpha level of .05 for all statistical analyses reported in this article. As is common in prospective memory research, prospective memory performance was measured in terms of the proportion of prospective memory targets to which the participant made a correct prospective memory response. If the tilde key was pressed at any time between the point at which one of the six prospective memory target words occurred and the point at which the next word was presented, this was counted as a correct prospective memory response. In Experiment 1, all participants pressed the tilde key alone rather than before or after the *Y* or *N* key.

Participants in the PMI group pressed the tilde key in response to a target word more often than did participants in the CMI group, $F(1, 62) = 5.05$, $MSE = 0.21$ (see Table 1). Trial type did not influence prospective memory accuracy, and trial type and group did not interact ($F\text{s} < 1$). The proportion of prospective memory target words recalled on the posttest questionnaire was not significantly affected by task importance, $t(62) = 0.12$. The mean number of targets recalled, out of six possible targets, was 4.77 ($SEM = 0.14$).

Multinomial Modeling Results

The frequencies of *Y* (color match), *N* (nonmatch), and prospective memory responses for the three different item types are presented in Appendix B, aggregated across participants and trials. From these frequencies, we obtained parameter estimates via maximum-likelihood estimation and goodness-of-fit index G^2, which is asymptotically chi-square distributed (Hu & Batchelder, 1994), for each instruction condition by using Excel Solver (Dodson, Prinzmetal, & Shimamura, 1998) and MBT (Hu, 1999). With $N = 1{,}984$ (32 participants × 62 test items), four degrees of freedom, and alpha = .05, the statistical power of these goodness-

Table 1
Experiments 1 and 2: Mean Proportion of Correct Prospective Memory Responses

Instruction condition	Experiment 1		Experiment 2: Target type — Same		Experiment 2: Target type — Different	
	M	*SE*	*M*	*SE*	*M*	*SE*
PMI	.65	.05	.87	.03	.78	.04
CMI	.46	.06	.76	.05	.67	.05

Note. PMI = emphasis on the prospective memory task; CMI = emphasis on the color-matching task.

of-fit tests was .96 to detect even a small effect size ($w = .1$; see J. Cohen, 1988).[4]

In the PMI and the CMI conditions, G^2s(4) were 1.61 and 2.09, respectively. Both values were smaller than the critical value of 9.49, indicating a good fit of the model to the data. That is, the assumptions made by our model, including the parameter constraints explained in the introduction, were met by the empirical data. When the additional constraint that $C_1 = C_2$ was imposed, the model no longer fit the data for either the PMI condition, $G^2(1) = 56.72$, or the CMI condition, $G^2(1) = 63.24$; the critical value of $G^2(1) = 3.84$. The fact that the model did not fit when C_1 and C_2 were set equal was not surprising, given the difference in baseline accuracy on match and nonmatch trials in the color-matching task.

The second step in the evaluation of the model is the experimental validity test. Do the parameter estimates, shown in Figure 2, adequately reflect the experimental manipulation? Significance tests revealed that the parameter estimates followed the predicted pattern: The estimate of preparatory processing, *P*, was significantly higher for the PMI group than it was for the CMI group, $G^2(1) = 18.54$. By contrast, there were no statistically significant group differences in the estimates of the contributions of memorial processes, *M*, $G^2(1) = 2.04$. No other tests of between-groups parameter differences yielded significant results. Thus, as expected, the manipulation of task importance affected the estimate of the *P* parameter independently of any other parameters.

One assumption of MPT models is that the observations are independent and identically distributed (Batchelder & Riefer, 1999). Our observations may violate this assumption, because we estimated parameter values based on data collapsed across participants (as is commonly done in multinomial modeling). Although MPT models have been shown to be rather robust with respect to this violation (Batchelder & Riefer, 1999; Erdfelder, in press; Riefer & Batchelder, 1991), we performed two additional sets of analyses to determine whether there were potential problems of heterogeneity in this experiment. The first set of analyses addressed concerns about possible differences between participants. The second addressed concerns about differences between items.

In the first set of analyses, participants within each experimental group (PMI and CMI) were randomly split into two subgroups of 16 participants each. The four-parameter model fit the data from each of the subgroups. The two subgroups were then compared to determine whether their parameter estimates differed. None of the parameters differed significantly between the subgroups in either the PMI or the CMI condition, all $G^2(1)\text{s} < 2.07$.

In the second set of analyses, the frequencies of *Y* (color match), *N* (nonmatch), and prospective memory responses for the three different item types were determined separately for the two different counterbalanced item lists. The estimates for each item list were based on 16 participants. The four-parameter model fit the data for each of the two item lists. The parameter estimates for both item lists were compared within each task-emphasis condition. There were no significant differences between item lists for either the PMI or the CMI condition, all $G^2(1)\text{s} < 1.96$. In sum, the results were robust across subgroups of participants and items.

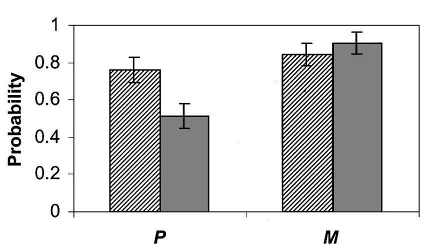

Figure 2. Multinomial model parameter estimates for Experiment 1. Striped bars indicate that instructions emphasized the importance of the prospective memory component of the task; solid bars indicate that instructions emphasized the importance of the color-matching component of the task. Error bars represent 95% confidence intervals. P = preparatory attentional processes; M = memory processes involved in discriminating between targets and nontargets.

Performance in the Ongoing Color-Matching Task

Given the PAM theory's proposal that the preparatory processes are resource demanding and that they play a fundamental role in prospective memory, one should expect to see a cost to ongoing activity associated with engaging in these processes, and the extent of the cost should be related to prospective memory performance (Smith, 2003). Evidence for this cost can be found by examining ongoing-task performance. The analysis of ongoing-task performance is also important for validating the model. If there is a cost to the ongoing task, then the model should reflect this in the *P* parameter.

We excluded the four trials immediately following each target trial (two trials following the final target), as well as the target trials, from the analyses of color-matching task performance in all four experiments in order to avoid finding an artifactual cost associated with response processes in the prospective memory condition. We also omitted the first four trials in each block. Accuracy and reaction time on accurate trials were examined for the remaining 30 nonprospective memory target trials in each block as a function of instruction group and trial type (match and nonmatch). As in previous studies (Smith, 2003; Marsh et al., 2003), the inclusion of the prospective memory task did not affect accuracy measures; thus, we focus on response time in this article.[5]

The baseline reaction times in Block 1 were significantly affected by neither trial type, $F(1, 79) = 2.54$, $MSE = 17{,}151$, nor group, and group and trial type did not interact ($F\text{s} < 1$). The mean baseline reaction time in Block 1 was 965 ms, $SEM = 27.00$.

[4] The power analyses reported in this article were performed with the GPOWER program by Erdfelder, Faul, and Buchner (1996).

[5] Statistical analyses for color-matching task accuracy were conducted on d' scores and on the proportion of correct trials. Across all four experiments, there were no significant effects or interactions involving the d' scores in either block of the color-matching task, and there were no significant effects on a difference score calculated by subtracting a participant's d' score in Block 1 from his or her d' score in Block 2. The proportion of correct responses was higher for nonmatch trials relative to match trials in all cases. This difference was significant in all four experiments, except for Experiment 3, in which there was a nonsignificant trend in favor of nonmatch trials, $p < .06$. The proportion of correct color-matching trials, which ranged from .84 to .91 for match trials and from .92 to .95 for nonmatch trials, was not significantly affected by any other manipulations.

Individual difference scores were computed by subtracting each participant's mean reaction time in Block 1 from their mean reaction time in Block 2. The mean difference scores are shown in Table 2.

Control group participants showed a practice effect: Reaction times were faster in Block 2 than in Block 1, $t(17) = 3.46$. Clearly, any cost found in the prospective memory groups could not be attributed to fatigue. The difference scores for both the PMI group, $t(31) = 9.48$, and the CMI group, $t(31) = 7.66$, were significantly different from zero, indicating a cost of performing the prospective memory task in both cases. As predicted, the higher prospective memory performance in the PMI group came at a greater cost for this group relative to the CMI group, $t(62) = 2.67$. Furthermore, there was a significant positive relationship between prospective memory performance and cost in both the PMI group ($r = .654$, $p < .01$) and the CMI group ($r = .650$, $p < .01$). This indicates that the slowing on the color-matching task is functionally related to prospective memory task performance. The color-matching task reaction-time results are consistent with the PAM theory's proposal that capacity-absorbing preparatory processes contribute to prospective memory performance.

Experiment 2

In Experiment 1, we took an initial step in validating the new multinomial model. In addition to demonstrating a good fit of the model to the data, in Experiment 1, we showed that the parameter for preparatory attentional processes, *P*, reflected the expected effect of a manipulation of the importance of the prospective memory task. Our objectives in Experiment 2 were twofold. First, because of the crucial importance of the preparatory-processes parameter, *P*, for PAM theory, we wanted to further corroborate the validity of this parameter. Second, we sought to evaluate the validity of the second important model parameter, namely the parameter, *M*, that is supposed to measure retrospective memory processes. To meet both objectives, we designed an experimental manipulation that should lead to a dissociation between preparatory and memory processes. That is, we manipulated the memory parameter in a predictable and opposite fashion from the change in the estimate for preparatory attentional processes. We achieved this by manipulating both task importance instructions (as in Experiment 1) and the similarity of the target words to the nontarget words. In Experiment 1, all of the target and nontarget words were unrelated. In Experiment 2, the semantic similarity of targets and nontargets was manipulated. Target and nontarget items either came from the same taxonomic categories (e.g., fruit) or from different taxonomic categories. We predicted that with increased semantic similarity of targets and nontargets, item recognition (i.e., discrimination between targets and nontargets) as measured by parameter *M* would decrease.

The presentation of multiple items from the same category can increase the processing of similarity among the items, but it does so at the expense of distinctive processing of each item (Hunt & Seta, 1984). The decrease in distinctive processing can reduce memory for the items (Hunt & McDaniel, 1993). By selecting multiple nontarget items that come from each of a limited number of categories, the similarity among the background items will be increased. If the targets also come from the same categories, the similarity among all items will be increased relative to when the targets come from different categories than the nontarget items, thereby decreasing the distinctiveness of the target items. This should decrease the discriminability of target and nontarget items. This would be consistent with research involving recognition memory tests. For example, in a recent study by Hunt (2003), participants were less able to distinguish between targets and nontargets when the targets and nontargets came from the same taxonomic category than they were when they came from different taxonomic categories (see also Bayen et al., 1996). Therefore, we predicted that the estimate of memory processes, *M*, would be lower for the condition in which the targets and nontargets come from the same categories than it would be for the condition in which the targets are selected from different categories than the nontargets.

Although the selection of targets from the same categories as the nontargets would decrease the *M* parameter, we predicted that the estimate of the contribution of preparatory processes would be larger in the same-category condition than it would be in the different-category condition. Participants in the different-category group would need to engage in fewer strategic preparatory processes because of the increased discriminability of the targets. We predicted that this decrease in preparatory processes should be reflected in the estimates of *P*. Moreover, in the same-category condition, the similarity between targets and nontargets should serve to remind the participants of the prospective memory task, thereby increasing the likelihood of preparatory processes. The presence of a cue can improve prospective memory performance (Einstein & McDaniel, 1990). Thus, we expected the presence of filler items that were highly related to the prospective memory targets to improve prospective memory performance, potentially offsetting any advantage of the distinctive targets.

In addition to the estimates for the memory and preparatory attentional processes, the experiment yielded response-time measures that provide an independent measure of cost associated with the preparatory processes. The predicted difference in the parameter *P* between the two conditions should be accompanied by a corresponding difference in the cost to the ongoing task. This translates into a prediction of a smaller difference in response times from the first block to the second block for the different-category condition than for the same-category condition.

In previous studies of prospective memory in which researchers used prospective memory target events that differ on some dimension from nontarget events, prospective memory performance improved. For instance, when targets were unfamiliar or low-frequency words and nontargets were familiar or high-frequency words, then prospective memory performance was better than it

Table 2
Experiment 1: Mean Difference Score for Reaction Time From First to Second Block of Color-Matching Trials (in Milliseconds)

Instruction condition	*M*	*SE*
PMI	579	61
CMI	371	48
Control	−100	29

Note. PMI = emphasis on the prospective memory task; CMI = emphasis on the color-matching task.

was when the target words were also familiar or high-frequency (Brandimonte & Passolunghi, 1994). Although previous researchers have manipulated the distinctiveness of prospective memory targets (Brandimonte & Passolunghi, 1994; Smith, 2003), none have done so using background items that are high in semantic similarity to the targets. We expected that the similarity between targets and nontargets would play an important role in the outcome of Experiment 2 and therefore did not expect to necessarily obtain the same performance results found in previous experiments using unrelated filler items. In our experiment, the predicted increase in preparatory processes in the same-category condition could outweigh any benefits from the distinctiveness of the targets in the different-category condition. Thus, we expected that the similarity of targets and nontargets would affect the parameter *M* but that the benefit for the distinctive targets in the different-category condition might be offset by increased preparatory processing in the same-category condition.

Finally, we predicted that the effects of task importance would be the same regardless of the relationship between target words and nontarget words. That is, the PMI group should have a higher estimate of preparatory processes than the CMI group, and the estimate of memory processes should not be influenced by task importance for both the same-category and different-category conditions. Furthermore, within an instructional group, the results for same category versus different category should follow the predictions outlined above.

To summarize, the memory-parameter estimate should be higher in the different-category condition than it is in the same-category condition. In contrast, the latter condition should see an increase in preparatory processes. The increase in preparatory processes in the same-category condition should be accompanied by a higher cost to the ongoing task relative to the different-category condition. The increase in preparatory processes might prevent the distinctive targets from having an advantage with respect to prospective memory performance. Therefore, no prediction was made concerning the effect of same category versus different category on prospective memory accuracy. The potential for opposite effects on the latent cognitive processes provides an important illustration of the advantages of using multinomial modeling: We can determine the effects of our manipulations on the attentional preparatory processes and memory processes, regardless of whether our manipulation affects the observed prospective memory accuracy. The effects of task importance, demonstrated in Experiment 1, should be replicated in both the same- and different-category groups.

Method

Design

Half of the 100 participants in this experiment were randomly assigned to the PMI condition, and half were assigned to the CMI group. Within each importance condition, 25 participants were randomly assigned to the same-category condition, and 25 were randomly assigned to the different-category condition. This between-subjects manipulation of the semantic similarity of the target and nontarget items is referred to as target type. As in Experiment 1, there was a within-subjects manipulation of trial type (color match vs. nonmatch). Thus, this experiment used a $2 \times 2 \times 2$ mixed factorial design.

Materials and Procedure

We selected 120 words from the Battig and Montague (1969) norms for this experiment in the following fashion. The following two groups of six categories each were selected: (a) a metal, a fish, a part of the human body, furniture, alcohol, and fruit and (b) precious stones, four-footed animals, an article of clothing, a part of a building, a substance for flavoring food, and a vegetable. Ten words were chosen for each category. Normally the 10 words were the top 10 instances, but some exceptions were made, such as removing spices and adding the 11th instance in the food-flavoring category, because the term *spices* could serve as a category label for some instances in the food-flavoring category. Similarly, *catfish*, *goldfish*, and *swordfish* were omitted from the fish category, and *grapefruit* was removed from the fruit list.

The fifth item in each category was selected to serve as a prospective memory target item. Thus there were two lists of six prospective memory targets: (a) *aluminum*, *perch*, *foot*, *desk*, *vodka*, and *peach* and (b) *pearl*, *lion*, *blouse*, *floor*, *vanilla*, and *potato*. Each list was used 50 times across the experiment. Targets occurred on Trials 10, 20, 30, 40, 50, and 60. Whether a trial was a match or a nonmatch trial was random. For most participants, half of the target trials were match trials, and half were nonmatch trials. Because of a programming error, some participants received one extra match and one fewer nonmatch trial or one extra nonmatch and one fewer match trial.

In the same-category condition, the nontarget items in the second block of color-matching trials, which included the embedded prospective memory task, came from the same categories as the target words. In the different-category condition, the nontarget items came from the other set of categories. The words that were not used in the second block were used in the first block of color-matching trials. The order of the nontarget words was random. There were a total of 54 nontarget trials, resulting in a total of 60 trials in each block of color-matching trials. The colors in the color-matching task were the same as those used in Experiment 1. The procedures were the same as in Experiment 1. No participants made both a prospective memory response and an ongoing-task response on any trials.

Results and Discussion

Prospective Memory Accuracy and Target Recall

Prospective memory accuracy, shown in Table 1, was not affected by trial type, and trial type did not interact with any other variables (Fs < 1). The likelihood of whether participants performed the prospective memory task was significantly influenced by target type, $F(1, 96) = 6.93$, $MSE = 0.05$, and by the task importance manipulation, $F(1, 96) = 4.70$, $MSE = 0.05$. The two variables did not interact ($F < 1$). Participants in the same-category condition were more likely than participants in the different-category condition to accurately respond to the prospective memory targets. Regardless of the type of target, the PMI instructions produced better prospective memory performance than the CMI instructions, providing a replication of Experiment 1. The number of prospective memory targets recalled on the posttest questionnaire was not affected by any of the manipulations (Fs $<$ 1). The mean number of targets recalled was 5.28 ($SEM = 0.10$) out of a possible 6 targets.

The increased prospective memory performance in the same-category condition relative to the different-category condition may appear counterintuitive, given previous findings showing a performance benefit for prospective memory targets that differ along some dimension from nontargets (e.g., Ellis & Milne, 1996). However, previous researchers who manipulated target distinctiveness did not use categorized lists. As we suggested in the introduction, it is possible that in the same-category condition, the participants were more likely to engage in preparatory processing, perhaps because there was an explicit recognition that the targets would be difficult to distinguish from the nontargets and that, therefore, a more strategic approach to the prospective memory task may have been used. It is also quite likely that the similarity between the targets and nontargets in the same-category condition would serve as a cue to remind the participant to engage in preparatory attentional processing, a hypothesis addressed in the following analyses. The interaction of the nature of the targets and the nature of the ongoing-task context in determining levels of prospective memory performance is consistent with previous research (e.g., Marsh, Hicks, & Hancock, 2000).

Multinomial Modeling Results

The response-category frequency counts aggregated over participants and trials are presented in Appendix C. Figures 3 and 4 show parameter estimates for each condition. With $N = 1{,}500$ (25 participants $\times$ 60 test items), four degrees of freedom, and alpha = .05, the statistical power of the goodness-of-fit tests was .89 for a small effect size ($w = .1$; see J. Cohen, 1988). $G^2(4)$ values for the four conditions were between 1.41 and 5.97 and thus were smaller than the critical value of 9.49, indicating a good fit of the model to the data. As in Experiment 1, parameter g was set to .10, and c was set to .50. The pattern replicates the effects of task importance seen in Experiment 1. Within each target-type group, estimates of P increased when the prospective memory task was emphasized relative to when the color-matching task was emphasized. This effect was significant with $G^2(1) = 4.34$ in the same-category condition and with $G^2(1) = 5.01$ in the different-category condition (greater than the critical value of 3.84). The memory parameter estimates remained unaffected by the manipulation of task importance. Unlike Experiment 1, there was a significant difference in the probability of detecting that the color of the word did not match, C_2, in the different-category condition, $G^2(1) = 4.23$. In the PMI condition, this probability was .90, whereas it was .85 in the CMI condition. This difference was somewhat surprising, given that the ability to detect that the color of the word did not match one from the preceding set increased for the group that had the emphasis on the prospective memory task. However, with a total of 20 significance tests in the two experiments and an alpha level of .05, we expected one test to turn out significant by chance alone. Despite the one anomaly that could be due to chance, the effects of task importance on the parameters of primary interest were as predicted and replicated the findings of Experiment 1.

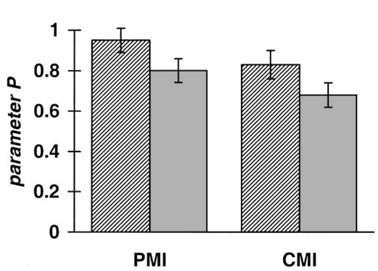

Figure 3. Estimates of parameter *P* (preparatory attentional processes) in Experiment 2. Striped bars represent the same-category condition; solid bars represent the different-category condition. Error bars represent 95% confidence intervals. PMI = instructions emphasized the importance of the prospective memory component of the task; CMI = instructions emphasized the importance of the color-matching component of the task.

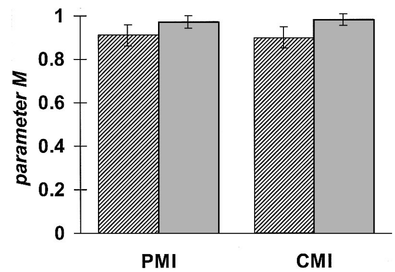

Figure 4. Estimates of parameter *M* (memory processes involved in discriminating between targets and nontargets) in Experiment 2. Striped bars represent the same-category condition; solid bars represent the different-category condition. Error bars represent 95% confidence intervals. PMI = instructions emphasized the importance of the prospective memory component of the task; CMI = instructions emphasized the importance of the color-matching component of the task.

The findings for the manipulation of target type were also as predicted. In both task-importance conditions, the parameter estimates, M, for the contribution of memory processes were larger for the different-category condition than they were for the same-category condition, $G^2(1) = 3.99$ for the PMI condition and $G^2(1) = 5.67$ for the CMI condition, indicating that the manipulation of semantic similarity between the target and nontarget events did impact the discriminability of targets and nontargets in memory. Also as predicted, participants in the same-category conditions were more likely to engage in preparatory processing (parameter P), regardless of task-importance instructions, $G^2(1) = 8.71$ for the PMI condition and $G^2(1) = 7.25$ for the CMI condition. Thus, the manipulation of target type had statistically significant and opposite effects on the two parameters of interest.

The multinomial modeling analyses provide important information that cannot be directly assessed when one is using more traditional empirical measures. The advantage regarding prospective memory performance for the targets in the same-category condition relative to the different-category condition can, with the help of multinomial modeling, be attributed to a difference in how participants perform the task. Namely, in the same-category condition, they are more likely to engage in preparatory attentional processes that help prospective memory performance.

Cost to the Ongoing Color-Matching Task

As in Experiment 1, mean response times were calculated for accurate color-matching trials. Response times in the first block of trials did not differ as a function of task importance, $F(1, 96) = 2.30$, $MSE = 76{,}278.84$, or as a function of target type ($F < 1$), and the two variables did not interact ($F < 1$). Therefore, a difference score was calculated for each participant by subtracting his or her mean response time in the first block from the mean response time in the second block.

The mean difference scores for all four groups, shown in Table 3 along with the test statistics for the single-sample comparisons, were significantly greater than zero. In other words, there was a

Table 3
Experiment 2: Mean Difference Score for Reaction Time From First to Second Block of Color-Matching Trials (in Milliseconds) and Value of Single-Sample Test Statistic

Instruction condition	Target type: Same — Difference score *M*	*SE*	*t*(24)	Target type: Different — Difference score *M*	*SE*	*t*(24)
PMI	641	47	13.71	498	47	10.64
CMI	333	60	5.53	249	54	4.62

Note. PMI = emphasis on the prospective memory task; CMI = emphasis on the color-matching task.

significant cost to performance on the ongoing task. The extent of the cost to the ongoing task was significantly affected by task importance, $F(1, 96) = 28.48$, $MSE = 68{,}187$, and by target type, $F(1, 96) = 4.73$, $MSE = 68{,}187$, but the two variables did not interact ($F < 1$). As in Experiment 1, the cost increased when the prospective memory task was emphasized. Also replicating Experiment 1, there was a significant positive correlation between cost to the ongoing-task and prospective memory performance for both the PMI condition ($r = .32$, $p < .03$) and the CMI condition ($r = .31$, $p < .03$). The response times on the color-matching task are consistent with our predictions and converge with the modeling results to support the argument that the similarity between the targets and nontargets in the same-category condition encourages more preparatory attentional processing.

Experiment 3

Our purpose in Experiment 3 was to further evaluate the model's validity by using an experimental manipulation that should selectively affect parameter *M*, which measures the retrospective memory processes, without affecting the *P* parameter. To this end, we manipulated the time available for encoding each prospective memory target. Half of the participants initially saw each target word for 5 s, and the other half for 20 s. In addition, recall of the target words was checked for the latter group, and the opportunity to restudy the target words was provided when needed. We predicted that participants in the 5-s group, the short encoding-time condition, would be less likely to perform the prospective memory task than participants in the long encoding-time condition. This prediction is based on previous research showing that prospective memory performance declines when encoding of the targets is made more difficult (Einstein et al., 1997).

Based on previous research showing better recognition performance following longer presentation time (e.g., Kellogg & Dare, 1989; Underwood, 1972), we predicted that the long-encoding group would be better than the short-encoding group at discriminating between targets and nontargets, reflected in a higher estimate of the *M* parameter for the long-encoding group in comparison to the short-encoding group. Presentation times at study should not affect the likelihood of engaging in preparatory attentional processes at the time of test, and it should not affect performance on the color-matching task. Thus, the *P*, C_1, and C_2 parameters should be statistically indistinguishable for the two experimental groups.

Method

Design

Half of the 52 participants in this experiment were randomly assigned to the short-encoding condition (5 s per target word) and half to the long-encoding condition (20 s per target word). There was also within-subject manipulation of trial type (match or nonmatch).

Materials and Procedure

The materials and procedures were the same as those used in Experiment 1, with the following exceptions. In Experiment 3, we did not include a delay between the prospective memory instructions and the start of the color-matching task because, based on results of Experiments 1 and 2, we did not believe that ceiling effects would occur on the prospective memory task. We also dropped the first block of color-matching trials because we were not concerned with providing evidence of a cost to ongoing activities, as this has been done in numerous experiments (Experiments 1 and 2 of this article; Burgess et al., 2001; Marsh et al., 2003; Smith, 2003; Smith & Hunt, 2004).

Participants read the color-matching task instructions and completed two color-matching trials. Following these two preliminary trials, the computer program paused and provided the participants with an opportunity to ask questions. When these questions were answered, the participant completed 16 practice trials for the color-matching task. After the practice trials, participants were given the prospective memory instructions, and they studied the prospective memory targets prior to starting the 62 trials of the color-matching task with the embedded prospective memory task.

The prospective memory instructions were changed to be neutral rather than emphasizing one task or the other, and the targets did not occur until after the instructions were completed. Participants read the instructions on the computer screen. The instructions said that in a few moments, they would learn some special words and that, if they saw these words in the color-matching task, they should try to remember to press the tilde key. They were instructed to ask if they had any questions about the task and to press any key to continue. When a key was pressed, the participant read that it was time to learn the target words and that each word would be shown on the screen, one at a time, for 5 s (in the short-encoding group) or 20 s (in the long-encoding group). The participant pressed a key to begin studying the target words. The six target words were shown, one at a time, on the screen for the appropriate amount of time in a random order. Following presentation of the prospective memory target words, participants in the short-encoding condition began the color-matching task. In the long-encoding condition, participants were asked to write the target words on a piece of paper. If they failed to recall all six targets, the words were presented again, one at a time, for 20 s each in a random order. No participant took more than two exposures to learn the target word to criterion. After participants in the long-encoding condition were able to recall all targets, the color-matching task was started.

Following the completion of the 62 color-matching trials, participants recalled the action and the prospective memory target words. All participants were able to recall the action.

Results and Discussion

Prospective Memory Accuracy and Target Recall

Participants in the long-encoding condition were more likely to press the tilde key when one of the target words occurred than were participants in the short-encoding condition, $t(50) = 2.09$. In the long-encoding condition, the mean proportion of prospective memory targets responded to correctly was .72, $SEM = .04$, compared with a mean proportion of .58, $SEM = .05$, in the short-encoding condition. Thus, as we predicted, increasing the difficulty of encoding the prospective memory targets by shortening the presentation time decreased the likelihood of performing the prospective memory task.

The mean number of prospective memory targets recalled, following completion of the color-matching task, out of 6 targets, was 2.96, $SEM = 0.33$, for the short-encoding condition and 5.35, $SEM = 0.21$, for the long-encoding condition. This difference was significant, $t(50) = 6.12$.

Multinomial Modeling Results

Appendix D contains the response-category frequencies aggregated over participants and trials for this experiment. As mentioned in the introduction, it is unusual in this sort of paradigm for participants to make both a color-matching task response and a prospective memory response on the same trial; however, 2 participants in this experiment did so on one or more trials. These responses were counted as prospective memory responses regardless of the color-matching response.

The same parameter constraints were set as in the analyses of the previous experiments. With $N = 1{,}612$ (26 participants $\times$ 62 test items), four degrees of freedom, and alpha = .05, the statistical power of the goodness-of-fit tests was .91 for a small effect size ($w = .1$; see J. Cohen, 1988). $G^2(4)$ values for the short- and long-encoding conditions were 5.35 and 0.80, respectively. These values are smaller than the critical value of 9.49, indicating a good fit of the model to the data.

The parameter estimates for P and M in each condition are shown in Figure 5. As we predicted, the encoding manipulation affected the M parameter. The ability to discriminate between the targets and nontargets was greater for the long-encoding condition than it was for the short-encoding condition, $G^2(1) = 13.93$ (larger than the critical value of 3.84). The encoding manipulation did not affect the probability of engaging in preparatory attentional processes measured by the parameter P, $G^2(1) = 0.31$. The parameters related to the ongoing task, C_1 and C_2, were also not affected by the encoding manipulation, $G^2(1) = 0.50$ and $G^2(1) = 3.81$, respectively. Thus, as predicted, the manipulation of encoding time

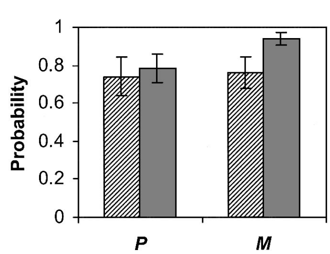

Figure 5. Estimates of parameters P (preparatory attentional processes) and M (memory processes for discriminating between targets and nontargets) in Experiment 3 as a function of prospective memory target-encoding time. Striped bars represent the short-encoding condition; solid bars represent the long-encoding condition. Error bars represent 95% confidence intervals.

affected the M parameter, although it left all other parameters unaffected.

Performance on the Ongoing Color-Matching Task

Evaluation of the color-matching task performance does not include a direct measure of cost, as there is no baseline condition for comparison. As mentioned earlier, we were not primarily concerned with demonstrating a cost to the ongoing task, as this had been done in Experiments 1 and 2, as well as in previous studies (e.g., Smith, 2003). However, performance on the ongoing task can be compared for the two groups. We did not expect the target-encoding manipulation to affect the extent to which participants engaged in preparatory processes. Therefore, the two groups should have similar resources available for performing the ongoing task, and performance on the color-matching task should be similar for the two groups.

Neither target-encoding time nor trial type had a significant effect on response times on the color-matching task, and the two variables did not interact (Fs < 1). The mean response time, in milliseconds, was 1,689.94, $SEM = 74.21$. In short, ongoing-task performance is consistent with our predictions that participants would have similar levels of resources available for the color-matching task in the two conditions. The results converge with the modeling: The encoding manipulation did not affect preparatory processing.

Summary

In Experiment 3, we provide additional validation of the model by demonstrating that the M parameter can be manipulated in a predictable fashion independent of any change in the P parameter. Additional support for the PAM theory comes from the finding that when the P parameter, thought to reflect the probability of engaging in resource demanding preparatory processes, does not change as a function of a particular manipulation, performance on the ongoing task is also unaffected.

Experiment 4

We had several interrelated goals in conducting Experiment 4. One was to replicate the results of Experiment 3, thus providing additional support for the validity of the multinomial model. We addressed this goal by using the same manipulation of target-encoding time in a new experiment with a different set of participants and modified procedures. A second goal was to use a different task format and a corresponding modified multinomial model in order to show that the application of multinomial models in prospective memory research is not restricted to the specific task format and the specific model used in Experiments 1–3.

A third and related goal was to address a possible concern that the preparatory processing postulated by PAM theory and measured by the multinomial model only occurs when participants must respond quickly, especially when they are not explicitly told that on any given trial, they can make a prospective memory response following the color-matching response. Thus, the color-matching task instructions in Experiment 4 made no mention about responding quickly, and participants were asked to respond to the color-matching task on every trial before responding to the prospective memory targets.

Everyday prospective memory tasks involve interrupting ongoing activities (Einstein & McDaniel, 1996), and we believe that instructions that do not require both responses capture this important aspect of prospective memory tasks more fully. Moreover, the task instructions used in Experiments 1–3 are in line with procedures used in numerous published studies (e.g., Burgess et al., 2001; A.-L. Cohen, Dixon, Lindsay, & Masson, 2003; A.-L. Cohen, West, & Craik, 2001; Einstein et al., 1997; Smith, 2003; West & Craik, 1999, 2001). Nonetheless, we wanted to rule out the possibility that our model would not fit a task in which both the prospective memory and ongoing-task responses are made. We changed the prospective memory instructions so that participants were explicitly instructed to respond to the color-matching task on every trial and then, on the appropriate target trials, to press the tilde key after making the color-matching response. This method was modeled after the instructions used by Marsh, Hicks, et al. (2002).

We predicted that the results would replicate the findings from Experiment 3. We expected that the encoding-duration manipulation would affect prospective memory accuracy and would affect retrospective recognition memory, as measured by the *M* parameter, but would not affect preparatory processes as measured by model parameter *P* and by response times on the ongoing task. We predicted that participants in the longer-encoding condition would have better prospective memory performance and a higher estimate of *M* than would participants in the short-encoding condition.

Method

Design

This experiment included the within-subject manipulation of trial type, match versus nonmatch, in addition to the between-subjects manipulation of long (20 s per word) versus short (5 s per word) prospective memory target-encoding condition. Twenty-five of the 51 participants were randomly assigned to the short encoding-duration condition, and 26 were randomly assigned to the long-encoding condition.

Materials and Procedure

Materials were the same as those used in Experiment 3. The procedures matched those in Experiment 3 with the following changes. The color-matching task instructions no longer said anything about responding to the color-matching task quickly. The displays seen by the participants on the color-matching task included an additional screen. Following the last color block on a trial, the word was displayed, and participants pressed the *Y* or the *N* key to make a color-matching task response, as in Experiments 1–3. In Experiments 1–3, the color-match response was followed by a blank screen and the automatic start of the next set of color blocks. Unlike in Experiments 1–3, now, after a participant pressed the *Y* or the *N* key, a screen was displayed with the request to “Please press the space bar to begin the next set.” Participants would make their color-matching responses and then press the space bar to begin seeing the next set of color blocks. This display was included to provide an opportunity to make the prospective memory response after the color-matching task response.

The prospective memory instructions were changed also. Participants were to remember to press the tilde key when the target words appeared as in Experiments 1–3. Unlike in Experiments 1–3, participants were explicitly told to make their color-match response first and then to make their prospective memory response after the color-match response, but before pressing the space bar to begin the next set of color blocks. The inclusion of the space bar screen and the instructions were based closely on the procedures used by Marsh, Hicks, et al. (2002).

All participants completed the same posttest questionnaire used in Experiments 1–3. One participant who failed to recall the intended action was dropped and replaced.

Results and Discussion

Prospective Memory Accuracy and Target Recall

Although they were instructed to make their prospective memory responses after making a color-match response, two participants in the short-encoding condition and four in the long-encoding condition made the prospective memory response instead of the ongoing-task response on one or more target trials. Therefore, the prospective memory accuracy and target-recall data were analyzed in two different ways. The first analysis included all participants. For participants who made a color-matching task response and a prospective memory response, the prospective memory response was recorded as such without regard to the color-matching task response. For those 6 participants who made a prospective memory response prior to the color-matching response, the prospective memory response was also considered to be correct. In the second analysis, the 6 participants who made at least one prospective memory response prior to making a color-matching task response were excluded from the analysis. Given that the results of the two analyses were the same, only the results of the first analysis are reported.

The mean proportion of prospective memory targets to which participants correctly responded was .80, $SEM = 0.06$, for the long-encoding condition and .62, $SEM = 0.05$, for the short-encoding condition. As predicted, the increased encoding time was followed by better prospective memory performance, $t(49) = 2.11$. There was also a significant effect of encoding condition on the number of target words recalled, out of six, on the posttest questionnaire, $t(49) = 4.49$. Participants in the short-encoding condition recalled a mean of only 2.60 targets, $SEM = 0.29$, in contrast to the mean of 4.54 targets, $SEM = 0.31$, recalled in the long-encoding condition.

Multinomial Modeling Results

The instructions used in Experiment 4, which requested that the ongoing-task response be made before the prospective memory response, allow the fitting of two different models to the data. The first model, which we call the three-response model, is the model that was used in the earlier Experiments 1–3. The second model, the four-response model, can now be fit to the data as well. In addition to the simple *Yes* and *No* responses that can be made on each trial, there are now two different prospective memory response possibilities. A prospective memory response could be made following a press of the *N* key or following a press of the *Y* key. Therefore, there are four responses possible on each trial: *Yes* only, *No* only, *Yes* followed by the prospective memory response, or *No* followed by the prospective memory response. The four-response model is shown in Figure 6 and is described in detail following the presentation of the results for the three-response model.

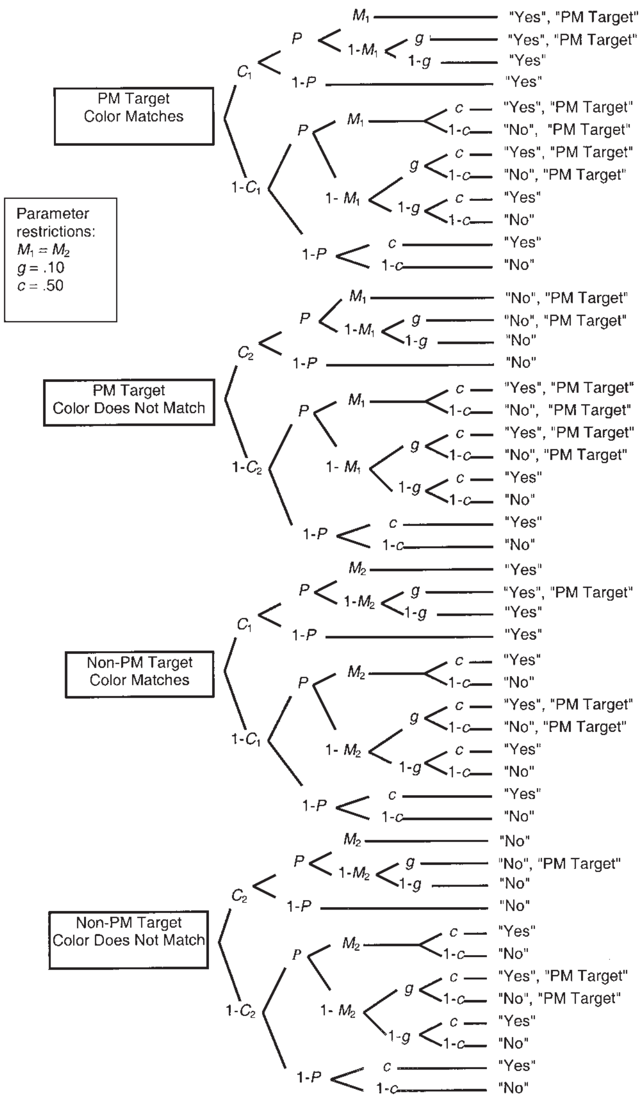

Figure 6. Multinomial model of event-based prospective memory for four response classifications. PM = prospective memory; C_1 = probability of detecting a color match; C_2 = probability of detecting that a color does not match; P = preparatory attentional processes; M_1 = probability of detecting a PM target; M_2 = probability of detecting that a word is not a PM target; g = probability of guessing that a word is a target; c = probability of guessing that a color matches.

Three-response model. When the color-matching task response was followed by the press of the tilde key, this was counted as a prospective memory response, regardless of whether the color-matching task response was *N* or *Y*. On those few cases in which the participant responded with the prospective memory response instead of the *Y* or *N* key, this was also counted as a prospective memory response. All 51 participants were included in this analysis. We also performed these analyses including only the 45 participants who responded to the color-matching task first on all trials (the participants included in the analysis of the four-response option model). The analyses yielded the same results in both cases.

The response-category frequencies aggregated across participants and trials for the three-response model analysis are presented in Appendix E. The same parameter constraints were applied as in Experiments 1–3. With $N = 1{,}550$ (25 participants $\times$ 62 test items), four degrees of freedom, and alpha = .05, the statistical power of the goodness-of-fit tests was .90 for a small effect size ($w = .1$; see J. Cohen, 1988). The $G^2(4)$ values for both the short and long encoding, 4.31 and 0.89, respectively, were smaller than the critical value of 9.49, indicating a good fit of the model to the data.

Figure 7 shows the parameter estimates for *P* and *M* for the long- and short-encoding conditions. As predicted, the encoding manipulation affected the *M* parameter but not the *P* parameter. Consistent with Experiment 3, the ability to discriminate between the targets and nontargets was better in the long-encoding condition than in the short-encoding condition, $G^2(1) = 27.23$ (larger than the critical value of 3.84). The probability of engaging in preparatory attentional processes, *P*, was not affected by the target-encoding manipulation, $G^2(1) = 0.18$. The parameters, C_1 and C_2, were not affected by the encoding manipulation. The findings from Experiment 3 were replicated here, supporting the validity of the model. Moreover, the model was found to fit these data even when the color-matching task did not require rapid responses and when the participants could make both a color-match response and a prospective memory response.

Four-response model. The four-response model is very similar to the three-response model, as can be seen in Figure 6. The only differences are reflected in the branches that lead to a prospective memory response. The prospective memory responses will occur under the same condition as in the three-response model, that is, when the participant engages in preparatory attentional processes and either correctly detects a target or guesses that an item is a target. However, rather than a prospective memory response alone, participants will now also make a color-match response. In the first and third trees, when a color match is correctly detected, which happens with probability C_1, the prospective memory response will be preceded by a *Yes* response on the color-matching task. When the color match is not detected, probability $1 - C_1$, the participant must guess whether the color matches. With probability *c*, the participant will guess that the color matches and will respond *Yes* on the color-match task followed by a prospective memory

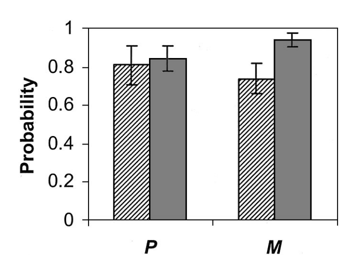

Figure 7. Parameter estimates for Experiment 4 based on the multinomial model for three response classifications. Parameters *P* (preparatory attentional processes) and *M* (memory processes for discriminating between targets and nontargets) as a function of prospective memory target-encoding time. Striped bars represent the short-encoding condition; solid bars represent the long-encoding condition. Error bars represent 95% confidence intervals.

response. If the participant guesses that the color does not match, probability $1 - c$, the prospective memory response will be preceded by a *No* response. In the second and fourth trees, the responses will be the same as it was in the first and third trees in those cases in which the participant fails to detect that a color does not match. That is, with probability *c*, the participant will guess that the color matches and respond *Yes* and then make the prospective memory response, and with probability $1 - c$, the participant will guess that the color does not match and respond *No* before making the prospective memory response. On the second and fourth trees, when a participant correctly detects that the color does not match, the prospective memory response will be preceded by a *No* response. The four-response model was found to be identifiable using the same methods shown in Appendix A for the three-response model.

Only those participants who made all of their prospective memory responses after making a color-matching task response were included in the following analyses. The raw data for these 47 participants are presented in Appendix F. Parameter constraints were the same as those chosen for the three-response model. With $N = 1{,}364$ (22 participants $\times$ 62 test items), eight degrees of freedom, and alpha = .05, the statistical power of the goodness-of-fit tests was .75 to detect a small effect size and greater than .99 to detect a medium effect size ($w = .1$ and .3, respectively; see J. Cohen, 1988). The $G^2(8)$ values for both the short and long encoding, 12.28 and 3.89, respectively, were smaller than the critical value of 15.51, indicating a good fit of the model to the data.

The estimates of *P* and *M* for the long- and short-encoding conditions, as determined by the four-response model, are shown in Figure 8. As with the estimates based on the three-response model, the encoding manipulation affected the *M* parameter estimate but not the *P* parameter estimate. Consistent with the findings from the three-response model in this experiment and in Experiment 3, the ability to discriminate between the targets and nontargets was better in the long-encoding condition than in the short-encoding condition, $G^2(1) = 18.28$. The parameter *P*, which measures the probability of engaging in preparatory attentional processes, was not affected by the target-encoding manipulation, $G^2(1) = 0.001$. As in the previous analyses using the three-response model, the encoding manipulation did not affect the parameters related to the ongoing task, C_1 and C_2.

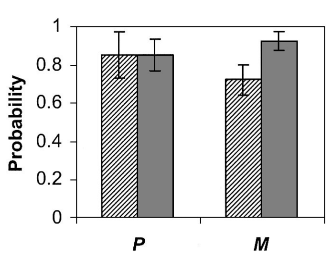

Figure 8. Parameter estimates for Experiment 4 based on the multinomial model for four response classifications. Parameters *P* (preparatory attentional processes) and *M* (memory processes for discriminating between targets and nontargets) as a function of prospective memory target-encoding time. Striped bars represent the short-encoding condition; solid bars represent the long-encoding condition. Error bars represent 95% confidence intervals.

Comparison of Experiments 3 and 4. Although one must exercise caution when making cross-experimental comparisons, the comparison of Experiments 3 and 4 can provide important information concerning the effects of the instructional changes on the parameter estimates. Furthermore, a comparison seems reasonable, given that the participants came from the same participant pool and were run in the same semester. Moreover, Experiment 4 was completed within 7 days of the completion of Experiment 3. Finally, an analysis of variance (ANOVA) for the vocabulary scores for the participants in the two experiments showed no effect of experiment, and experiment and encoding condition did not interact (*F*s < 1). Encoding condition also had no effect on vocabulary scores, $F(1, 99) = 1.08$, $MSE = 0.02$.

In Experiment 4, we addressed the possibility that, in the absence of a time demand and with a task that requested that a prospective memory response be made following the color-matching response, a model that includes the preparatory processing parameter would no longer fit the data. We refer to the possibility that the model only fit in Experiments 1–3 because of the specific task instructions as the *task-demands explanation*. In contrast to the prediction based on the task-demands explanation, the model provided a good fit to the data in Experiment 4. In addition to predicting that the model might not fit data using the alternative procedures of Experiment 4, the logic of the task-demands explanation can be extended to predict that the relaxing of the task demands in Experiment 4 would, at the very least, decrease the demands for preparatory processing, leading to a reduction in the estimate of the *P* parameter in Experiment 4 relative to Experiment 3. The estimates of *P* were compared for the short-encoding conditions and the long-encoding conditions separately. In neither the short-encoding condition, $G^2(1) = 1.06$, nor the long-encoding condition, $G^2(1) = 1.48$, did the estimate of *P* change from Experiment 3 to Experiment 4. The estimates of *M* were also unchanged from one experiment to the next. Furthermore, the proportion of correct prospective memory responses also remained unchanged from Experiment 3 to Experiment 4. An ANOVA, with the two independent variables of experiment and encoding condition, produced a main effect of target-encoding duration, $F(1, 99) = 10.49$, $MSE = 0.07$, but no effect of experiment, $F(1, 99) = 1.20$, $MSE = 0.07$, and no interaction of the two variables, $F < 1$.

One might wish to argue that the lack of a change in *P* from Experiment 3 to Experiment 4 indicates that change in instructions did not really change the way participants approached the task. However, the comparison of color-matching task performance across Experiments 3 and 4, discussed in the *Performance on the Ongoing Color-Matching Task* section, argues against this.

Performance on the Ongoing Color-Matching Task

As in Experiment 3, we expected no effect of group on color-match response times, given that we predicted that the encoding manipulation would not affect the extent of preparatory processing, and that, therefore, the two groups should have similar levels of resources available for performing the ongoing task. The analyses reported included all participants. We found the same results when including only those 45 participants who responded with a color-matching response on every trial.

Consistent with the findings in Experiments 1–3, response times on the color-matching task did not differ as a function of trial type, $F < 1$. Encoding condition also did not affect response times, $F(1, 49) = 2.94$, $MSE = 432{,}333.98$, and the two variables did not interact, $F < 1$. The mean response time on accurate color-matching trials was 2,104.91 ms, $SEM = 93.84$. The color-matching task results are consistent with Experiments 1–3 reported here and also with the prediction that the encoding manipulation would not affect color-match task performance.

We compared reaction times on the ongoing task in Experiment 4 with reaction time to the ongoing task in Experiment 3. The change in instructions led to significantly longer response times in Experiment 4 than in Experiment 3 on the color-matching task, $F(1, 99) = 12.16$, $MSE = 359{,}061.91$. Thus, the change in instructions influenced how the color-matching task was performed but did not change prospective memory task performance. In short, the results do not support the task-demands explanation for why our model, which includes an estimate of preparatory processing, provided a good fit to the data in these experiments.

In summary, the multinomial modeling results for Experiment 4 provide a replication of Experiment 3, thereby further supporting the validity of the model. In addition, Experiment 4 provides evidence that the fit of the model in Experiments 1–3 was not an artifact of using an ongoing task that demanded rapid responses.

General Discussion

In this study, we introduce a formal model of event-based prospective memory that provides a means for distinguishing between preparatory processes that are engaged prior to the target event and retrospective memory processes that are engaged once a target occurs. In addition to introducing the use of formal mathematical models in the area of prospective memory research, this research served to validate the new model. In Experiment 1, a manipulation of task importance increased the estimate of the contribution of preparatory attentional processes without increasing the estimate of memory processes. This pattern was replicated in two different conditions in Experiment 2. In Experiment 2, participants either had distinctive prospective memory targets, which came from different taxonomic categories than the background items, or nondistinctive prospective memory targets, which came from the same taxonomic categories as the background items. In both target-type conditions, the manipulation of task importance affected only the parameter for preparatory processing. The manipulation of target type, regardless of task importance, had the predicted opposite effects on the parameters for preparatory attentional processes and memory processes. Finally, in Experiments 3 and 4, we manipulated the amount of time participants had available to encode each prospective memory target. When participants had more time to learn the target words, their prospective memory performance increased. Modeling revealed that the manipulation of target-encoding time affected memory processes only.

Experiment 2 provides a particularly clear illustration of the benefits of mathematical modeling. In this experiment, we manipulated the distinctiveness of the prospective memory targets by changing the semantic similarity between the target words and the nontarget words. Multinomial modeling revealed that, as expected, the discriminability of targets and nontargets was higher in the distinctive (i.e., different-category) condition. However, preparatory attentional processes were more likely in the nondistinctive (i.e., same-category) condition. The increase in preparatory attentional processes in this condition outweighed the memory benefits of the distinctive targets, and prospective memory performance was higher in the nondistinctive condition. If, in the absence of a model, we had looked at the empirical prospective memory performance measure only, it would be difficult to know whether we had failed to manipulate target distinctiveness, or whether we had a more complicated effect of the semantic-similarity manipulation. Multinomial modeling of the data from this experiment enabled us to measure the dissociation of preparatory attentional and memory processes.

To further illustrate this point, the following example provides a demonstration of how the model could predict equivalent proportions of correct prospective memory responses despite variations in model parameters. For the four-parameter model with three response classifications, the model equation for the predicted proportion of correct prospective memory (PM) responses is

$$\text{PM hit rate} = C_1\,P\,M + C_1\,P\,(1-M)\,g + (1-C_1)\,P\,M + (1-C_1)\,P\,(1-M)\,g + C_2\,P\,M + C_2\,P\,(1-M)\,g + (1-C_2)\,P\,M + (1-C_2)\,P\,(1-M)\,g \quad (1).$$

We can substitute parameter values to determine the predicted proportion of prospective memory hits. For instance, assume one condition that encourages preparatory processing and that includes targets and nontargets that are not particularly easy to discriminate. Assume for this condition that $P = .80$ and $M = .50$. Contrast this with a condition in which engaging in preparatory processing is less likely, perhaps because of a very demanding ongoing task, but in which the discrimination of targets and nontargets is easier. Assume that in this condition, $P = .60$ and $M = .70$. To make the example straightforward, assume that C_1 and C_2 are the same for both groups, .70 and .85, respectively. Further assume that $g = .10$ for both conditions as we did for our experiments. Substituting these parameter values into Equation 1 results in a proportion of prospective memory hits of .88 in the first case and .876 in the second case. Thus, opposite patterns of latent cognitive processes may underlie almost identical levels of performance as measured by an empirical performance measure. Only a formal model can reveal the latent processes.

The model is applicable to prospective memory paradigms that involve ongoing tasks with two trial types, such as match versus nonmatch in the experiments reported here, and two response possibilities, such as *Yes* or *No,* for the ongoing task in these experiments. McDaniel, Robinson-Riegler, and Einstein (1998) embedded an event-based prospective memory task in an ongoing sentence-verification task in which sentences were either true or false and in which participants had to respond *Yes* or *No* concerning the veracity of the sentences. We have successfully applied the model to a data set involving this sentence-verification task (Smith & Bayen, 2002).

In Experiments 1 and 2, we had an independent index of attentional processes in the form of a cost measure, that is, a reaction-time difference score. The advantage of using the model over the reaction-time difference score is that the model can be applied to situations in which reaction times are not sufficiently sensitive, such as self-paced tasks in which the speed of responses is not emphasized. We have successfully applied the model in such cases (Experiment 4 in this article; Smith & Bayen, 2002).

Using the Model to Contrast Alternative Frameworks

The multinomial model presented in this article was driven by a particular theoretical position. The PAM theory is similar to the multiprocess framework proposed by McDaniel and Einstein (2000). The primary difference between the multiprocess view and PAM is that according to the multiprocess view, some event-based tasks may be accomplished automatically, whereas PAM proposes that resource-demanding preparatory attentional processes are always required for successful prospective memory performance.

We can contrast these two perspectives here by setting values for the preparatory attentional processes parameter to reflect an automatic-retrieval account of prospective memory and testing the goodness of fit of this model to the data. According to the multiprocess framework, there "may be a handful of relatively automatic processes that can support prospective-memory retrieval" (McDaniel & Einstein, 2000, p. S131). The multiprocess view encompasses a number of different explanations that have been proposed for how intentions might be retrieved without relying on preparatory attentional processes. In two of these frameworks, the automatic associative activation framework (Guynn et al., 2001) and the simple activation framework (Einstein & McDaniel, 1996), the occurrence of a target will automatically bring to mind the action, given a sufficiently well encoded target–action relationship. According to the notice + search view (Einstein & McDaniel, 1996), the occurrence of a target event will produce a feeling of familiarity, or a noticing, which in turn will lead to the search of memory for why the target is familiar. The important point is that none of the "automatic" explanations that are encompassed by the multiprocess view require the engagement of resource-demanding preparatory processes prior to the occurrence of the target event.

Because preparatory processes are not included in these explanations, the *P* parameter would play no role in successful prospective memory performance and should therefore be set to a constant value. This is similar to the ideas expressed in Kliegel et al. (2001) when they predict that increased monitoring that might accompany an important event-based prospective memory task will not improve performance. Because the retrieval of the intention is automatic, it does not matter what sort of monitoring happens before the target occurs. Automatic means that the processes are immune to any controlled attentional factors. We set the *P* parameter to one to reflect this case. It may seem paradoxical to set $P = 1$ rather than $P = 0$ to test the automatic-retrieval type of explanation, given that $P = 1$ when there is perfect probability of monitoring. Setting the preparatory attentional processes parameter to one, however, can also represent the case in which no monitoring is required to achieve prospective memory performance, because the intention pops into mind automatically, regardless of whether monitoring occurs. In this case, as can be seen in Figure 1, any variation in performance depends on variations in the memory processes. When parameter *P* is set to zero, no prospective memory performance can be achieved according to the model. Thus, to evaluate the automatic-activation types of theories, the fit of the multinomial model was evaluated when $P = 1$.

When the model was applied with $P = 1$, it did not fit; that is, the values of $G^2(5)$ exceeded the critical value of 11.07 for each of the data sets in the experiments presented here, with one exception. Only in the case of prospective memory important instructions and same-category targets in Experiment 2 did the model fit the data. This is the situation in which the highest estimate of $P = .95$ was obtained when *P* was allowed to vary, so it is not that surprising that when *P* is set to one, the model still fits the data. The combination of instructions and similarity between the targets and nontargets in this case encourages extensive monitoring, and, therefore, the model with $P = 1$ would be expected to fit. Here, $P = 1$ reflects a perfect probability of engaging in preparatory processing. This is also a case in which the multiprocess view would agree that the prospective memory task is not automatic, because the target salience was reduced because of the similarity between the targets and nontargets, and more strategic processing could have been encouraged by the instructions. In other words, when the model is set to reflect an automatic-activation type of explanation, the model no longer provides a good fit to the data, except in a case in which the task should not be automatic. The case which seems to best, although not perfectly,[6] meet the multiprocess view's requirements for automaticity is the case in which the background task was emphasized and the targets were arguably most salient: the CMI condition of Experiment 2, in which the targets came from different categories than the nontarget items. The automaticity model does not fit the data in this case. Thus, the evidence suggests that the event-based prospective memory tasks in these experiments were not being accomplished through automatic retrieval of the action. One might argue that the event-based prospective memory tasks in these experiments were ones that cannot be accomplished automatically, and there may still be other situations in which the event-based prospective memory tasks are automatic. This is certainly possible. Additional research and clarification of the theories is needed in order to determine when, if ever, the automatic retrieval of an intention does occur. We hope that the introduction of formal mathematical modeling to the area of prospective memory research will encourage and facilitate the development of more fully evolved theoretical positions.

Finally, our model may ultimately be compatible with other explanations of prospective memory. If the multiprocess framework's "largely automatic" tasks (McDaniel & Einstein, 2000) refer to cases

[6] This case may not strictly fit the restrictions for an automatic task proposed by the multiprocess view because, according to this view, in order for a target event to trigger an automatic retrieval of the intention, the ongoing task must focus on event features that are specified in the prospective memory intention (McDaniel & Einstein, 2000). Although our ongoing color-matching task does focus on the target events, the demands of the color-matching task require processing of the color of the word, not the meaning, and the meaning (or specific word) is relevant to the prospective memory task. Word reading has been shown to be automatic when the task is to name the color the word is printed in when the entire word is printed in color, which is similar to the task used here (e.g., Besner, Stoltz, & Boutilier, 1997; Kahneman & Henik, 1981; Kahneman & Treisman, 1984), and therefore, we argue that the color-matching task does involve processing of the relevant features, in this case the meaning, of the target event. However, proponents of the multiprocess view might not agree that the color-matching task involves processing of the relevant features of the target event.

that do require resources, then the framework and PAM theory may be compatible. PAM theory may also share common threads with explanations in which a supervisory attentional system plays a role (e.g., Ellis, 1996). The precise role of automatic versus nonautomatic processes in these explanations is not entirely clear, making it difficult to evaluate the relationship of the PAM theory to these frameworks. The use of MPT models in the area of prospective memory research might help to clarify these issues.

References

Andrzejewski, S. J., Moore, C. M., Corvette, M., & Herrmann, D. (1991). Prospective memory skill. *Bulletin of the Psychonomic Society, 29,* 304–306.

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97,* 548–564.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57–86.

Battig, W. F., & Montague, W. E. (1969). Category norms of verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology, 80,* 1–46.

Bayen, U. J., & Murnane, K. (1996). Aging and the use of perceptual and temporal information in source memory tasks. *Psychology and Aging, 11,* 293–303.

Bayen, U. J., Murnane, K., & Erdfelder, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 197–215.

Bayen, U. J., Nakamura, G. V., Dupuis, S. E., & Yang, C.-L. (2000). The use of schematic knowledge about sources in source monitoring. *Memory & Cognition, 28,* 480–500.

Besner, D., Stoltz, J. A., & Boutilier, C. (1997). The Stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review, 4,* 221–225.

Brandimonte, M. A., & Passolunghi, M. C. (1994). The effect of cue-familiarity, cue distinctiveness, and retention interval on prospective remembering. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 47*(A), 565–587.

Buchner, A., Erdfelder, E., & Vaterrodt-Plünnecke, B. (1995). Toward unbiased measurement of conscious and unconscious memory processes within the process dissociation framework. *Journal of Experimental Psychology: General, 124,* 137–160.

Burgess, P. W., Quayle, A., & Frith, C. D. (2001). Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia, 29,* 545–555.

Cohen, A.-L., Dixon, R. A., Lindsay, D. S., & Masson, M. E. J. (2003). The effects of perceptual distinctiveness on the prospective and retrospective components of prospective memory in young and old adults. *Canadian Journal of Experimental Psychology, 57,* 274–289.

Cohen, A.-L., West, R., & Craik, F. I. M. (2001). Modulation of the prospective and retrospective components of memory for intentions in younger and older adults. *Aging, Neuropsychology, and Cognition, 8,* 1–13.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Dodson, C. S., Prinzmetal, W., & Shimamura, A. P. (1998). Using Excel to estimate parameters from observed data: An example from source memory data. *Behavior Research Methods, Instruments, & Computers, 30,* 517–526.

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 717–726.

Einstein, G. O., & McDaniel, M. A. (1996). Retrieval processes in prospective memory: Theoretical approaches and some new empirical findings. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 115–142). Mahwah, NJ: Erlbaum.

Einstein, G. O., Smith, R. E., McDaniel, M. A., & Shaw, P. (1997). Aging and prospective memory: The influence of increased task demands at encoding and retrieval. *Psychology and Aging, 12,* 479–488.

Ellis, J. A. (1996). Prospective memory or the realization of delayed intentions: A conceptual framework for research. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 1–22). Mahwah, NJ: Erlbaum.

Ellis, J., & Milne, A. (1996). Retrieval cue specificity and the realization of delayed intentions. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49*(A), 862–887.

Erdfelder, E. (in press). *Multinomiale Modelle in der kognitiven Psychologie: Eine Einführung* [Multinomial models in cognitive psychology: An introduction]. Cologne, Germany: Kölner Studienverlag.

Erdfelder, E., & Buchner, A. (1998). Decomposing the hindsight bias: A multinomial processing tree model for separating recollection and reconstruction in hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 387–414.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28,* 1–11.

Goschke, T., & Kuhl, J. (1996). Remembering what to do: Explicit and implicit memory for intentions. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 53–92). Mahwah, NJ: Erlbaum.

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

Hu, X. (1999). Multinomial processing tree models: An implementation. *Behavior Research Methods, Instruments, & Computers, 31,* 689–695.

Hu, X., & Batchelder, W. H. (1994). The statistical analysis of general processing tree models with the EM algorithm. *Psychometrika, 59,* 21–47.

Hunt, R. R. (2003). Two contributions of distinctive processing to accurate memory. *Journal of Memory and Language, 48,* 811–825.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language, 32,* 421–445.

Hunt, R. R., & Seta, C. E. (1984). Category size effects in recall: The roles of relational and individual item information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 454–464.

Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 181–211). Hillsdale, NJ: Erlbaum.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29–61). New York: Academic Press.

Kellogg, R. T., & Dare, R. S. (1989). Explicit memory for unattended information. *Bulletin of the Psychonomic Society, 27,* 409–412.

Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the "Who said what?" paradigm. *Journal of Personality and Social Psychology, 75,* 155–178.

Kliegel, M. K., Martin, M., McDaniel, M. A., & Einstein, G. O. (2001). Varying the importance of a prospective memory task: Differential effects across time- and event-based prospective memory. *Memory, 9,* 1–11.

Krishnan, H. S., & Shapiro, S. (1999). Prospective and retrospective memory for intentions: A two-component approach. *Journal of Consumer Psychology, 8,* 141–166.

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Kvavilashvili, L. (1987). Remembering intention as a distinct form of memory. *British Journal of Psychology, 78,* 507–518.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and "nonparametric" indexes. *Psychological Bulletin, 107,* 401–413.

Marsh, R. L., Hancock, T. W., & Hicks, J. L. (2002). The demands of an ongoing activity influence the success of event-based prospective memory. *Psychonomic Bulletin & Review, 9,* 604–610.

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

Marsh, R. L., Hicks, J. L., & Hancock, T. W. (2000). On the interaction of ongoing cognitive activity and the nature of an event-based intention. *Applied Cognitive Psychology, 14,* S29–S42.

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.

Maylor, E. A., Smith, G., Sala, S. D., & Logie, R. H. (2002). Prospective and retrospective memory in normal aging and dementia: An experimental study. *Memory & Cognition, 30,* 871–884.

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

McGann, D., Ellis, J., & Milne, A. (2002). Conceptual and perceptual processing in prospective remembering: Differential influences of attentional resources. *Memory & Cognition, 30,* 1021–1032.

McGann, D., Ellis, J., & Milne, A. (2003). Conceptual and perceptual processing in prospective remembering. *European Journal of Cognitive Psychology, 15,* 19–41.

Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 669–682). New York: Academic Press.

Ratcliff, R., Sheu, C. F., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review, 99,* 518–535.

Reese, C. M., & Cherry, K. E. (2002). The effects of age, ability, and memory monitoring on prospective memory task performance. *Aging, Neuropsychology, and Cognition, 9,* 98–113.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review, 95,* 318–339.

Riefer, D. M., & Batchelder, W. H. (1991). Statistical inference for multinomial processing tree models. In J.-P. Doignon & J.-C. Falmagne (Eds.), *Mathematical psychology: Current developments* (pp. 313–335). New York: Springer Publishing.

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27,* 763–776.

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. (2002, November). *Investigating the role of working memory capacity in prospective memory.* Poster session presented at the annual meeting of the Psychonomic Society, Kansas City, MO.

Smith, R. E., & Hunt, R. R. (2004). *Retrieving delayed intentions to salient target events requires capacity.* Manuscript submitted for publication.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34–50.

Spaniol, J., & Bayen, U. J. (2002). When is schematic knowledge used in source monitoring? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 631–651.

Swets, J. A. (1986). Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. *Psychological Bulletin, 99,* 100–117.

Underwood, B. J. (1972). Word recognition memory and frequency information. *Journal of Experimental Psychology, 94,* 276–283.

Van Zandt, T. (2000). ROC curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 582–600.

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

West, R., & Craik, F. I. M. (2001). Influences on the efficiency of prospective memory in younger and older adults. *Psychology and Aging, 16,* 682–696.

Yu, J., & Bellezza, F. S. (2000). Process dissociation as source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1518–1533.

Appendix A

Seven-Parameter Model: Not Identifiable

A necessary condition for the global identifiability of a model is that the number of dimensions of the parameter space not be greater than the number of dimensions of the space of possible response-category probabilities (Erdfelder, in press). In other words, the model may not have more free parameters than there are independent model equations. We now show that the seven-parameter model as depicted in Figure 1 does not meet this condition and is thus not globally identifiable.

Table A1 shows the response-category probability designations for the various item and response types. Each p_{ij} is the probability of reponse j to an item of type i.

The model as illustrated in Figure 1 has seven parameters in the parameter space:

$$\Omega_7 = [(C1, C2, P, M1, M2, g, c)]. \quad \text{(A1)}$$

The response probabilities in Table A1 can be expressed as equations, created by summing branch probabilities for the tree in Figure 1 for each response:

$$p_{11} = C_1P(1 - M_1)(1 - g) + C_1(1 - P) + (1 - C_1)P(1 - M_1)(1 - g)c + (1 - C_1)(1 - P)c \quad \text{(A2)}$$

$$p_{12} = (1 - C_1)P(1 - M_1)(1 - g)(1 - c) + (1 - C_1)(1 - P)(1 - c) \quad \text{(A3)}$$

$$p_{13} = C_1PM_1 + C_1P(1 - M_1)g + (1 - C_1)PM_1 + (1 - C_1)P(1 - M_1)g \quad \text{(A4)}$$

$$p_{21} = (1 - C_2)P(1 - M_1)(1 - g)c + (1 - C_2)(1 - P)c \quad \text{(A5)}$$

$$p_{22} = C_2P(1 - M_1)(1 - g) + C_2(1 - P) + (1 - C_2)P(1 - M_1)(1 - g)(1 - c) + (1 - C_2)(1 - P)(1 - c) \quad \text{(A6)}$$

$$p_{23} = C_2PM_1 + C_2P(1 - M_1)g + (1 - C_2)PM_1 + (1 - C_2)P(1 - M_1)g \quad \text{(A7)}$$

$$p_{31} = C_1PM_2 + C_1P(1 - M_1)(1 - g) + C_1(1 - P) + (1 - C_1)PM_2c + (1 - C_1)P(1 - M_1)(1 - g)c + (1 - C1)(1 - P)c \quad \text{(A8)}$$

$$p_{32} = (1 - C_1)PM_2(1 - c) + (1 - C_1)P(1 - M_2)(1 - g)(1 - c) + (1 - C_1)(1 - P)(1 - c) \quad \text{(A9)}$$

$$p_{33} = C_1P(1 - M_2)g + (1 - C_1)P(1 - M_2)g \quad \text{(A10)}$$

Table A1

Designation of Response Category Probabilities

Item type	Response type: Y	N	PM
Target, match	p_{11}	p_{12}	p_{13}
Target, nonmatch	p_{21}	p_{22}	p_{23}
Nontarget, match	p_{31}	p_{32}	p_{33}
Nontarget, nonmatch	p_{41}	p_{42}	p_{43}

Note. Y = color match; N = nonmatch; PM = prospective memory.

$$p_{41} = (1 - C_2)PM_2c + (1 - C_2)P(1 - M_2)(1 - g)c + (1 - C_2)(1 - P)c \quad \text{(A11)}$$

$$p_{42} = C_2PM_2 + C_2P(1 - M_2)(1 - g) + C_2(1 - P) + (1 - C_2)PM_2(1 - c) + (1 - C_2)P(1 - M_2)(1 - g)(1 - c) + (1 - C_2)(1 - P)(1 - c) \quad \text{(A12)}$$

$$p_{43} = C_2P(1 - M_2)g + (1 - C_2)P(1 - M_2)g. \quad \text{(A13)}$$

Within each item category, response-category probabilities add up to 1. Therefore, one response-category probability can be expressed in terms of the other two response probabilities for that item category. For instance, $p_{11} = 1 - p_{12} - p_{13}$. This restricts the number of independent equations, and therefore the degrees of freedom, to 8. For purposes of the following discussion, we designate Equations A2, A6, A8, and A12 as being expressed in terms of the other two response-category probabilities within their respective item categories. Equations A3, A4, A5, A7, A9, A10, A11, and A13 remain as potentially independent equations. However, we now demonstrate that four of these equations can also be expressed as a function of other response probabilities, therefore further decreasing the number of independent equations.

Equation A4 can be rewritten and simplified in the following way:

$$p_{13} = C_1PM_1 + C_1P(1 - M_1)g + PM_1 - C_1PM_1 + P(1 - M_1)g - C_1P(1 - M_1)g$$

$$p_{13} = PM_1 + P(1 - M_1)g \quad \text{(A4)}$$

$$p_{13} = PM_1 + Pg - PM_1g. \quad \text{(A14)}$$

In the same manner, Equation A7 can be reduced, showing that $p_{13} = p_{23}$. Similarly, Equations A10 and A13 show that $p_{33} = p_{43}$:

$$p_{33} = p_{43} = Pg - PM_2g. \quad \text{(A15)}$$

The equality of these response-category probabilities reduces the number of independent equations by two, leaving only six equations. We now demonstrate that two additional response-category probabilities can be expressed in terms of other response-category probabilities. Start by simplifying Equation A3:

$$p_{12} = (1 - C_1)P(1 - M_1)(1 - g)(1 - c) + (1 - C_1)(1 - P)(1 - c) \quad \text{(A3)}$$

$$p_{12} = (1 - C_1)(1 - c)[1 - (PM_1 + Pg - PM_1g)]. \quad \text{(A16)}$$

Using Equation A14, we can substitute p_{13} for $(PM_1 + Pg - PM_1g)$ in Equation A16:

$$p_{12} = (1 - C_1)(1 - c)(1 - p_{13}). \quad \text{(A17)}$$

Equation A17 can be rewritten as follows:

$$p_{12}/(1 - p_{13}) = (1 - C_1)(1 - c). \quad \text{(A18)}$$

In a similar fashion, Equation A9 can be rewritten:

$$p_{32} = (1 - C_1)(1 - c)[1 - (Pg - PM_2g)]. \quad \text{(A19)}$$

Using Equation A15, we can substitute p_{33} for $Pg - PM_2g$:

$$p_{32} = (1 - C_1)(1 - c)(1 - p_{33}). \quad \text{(A20)}$$

(*Appendixes continue*)

Equation A20 can be expressed as

$$p_{32}/(1 - p_{33}) = (1 - C_1)(1 - c). \quad \text{(A21)}$$

On the basis of Equation A18, we substitute for $(1 - C_1)(1 - c)$ in Equation A21:

$$p_{32}/(1 - p_{33}) = p_{12}/(1 - p_{13}). \quad \text{(A22)}$$

Equation A22 can be solved to show that p_{32} can be expressed as a function of p_{12}, p_{13}, and p_{33}:

$$p_{32} = [p_{12}/(1 - p_{13})](1 - p_{33}). \quad \text{(A23)}$$

In a similar way, by rewriting Equations A5 and A11 and making substitutions based on Equations A14 and A15, we can express p_{21} as a function of p_{41}, p_{13}, and p_{33}:

$$p_{21} = (1 - C_2)c(1 - p_{13})$$

$$p_{41} = (1 - C_2)c(1 - p_{33}) \quad \text{(A24)}$$

$$p_{21}/(1 - p_{13}) = (1 - C_2)c = p_{41}/(1 - p_{33})$$

$$p_{21} = [p_{41}/(1 - p_{33})](1 - p_{13}). \quad \text{(A25)}$$

Equations A23 and A25 demonstrate that the two additional response-category probabilities can be expressed in terms of other category probabilities, further reducing the number of independent equations to four. Thus, the model is not identifiable when there are more than four free parameters.

Demonstrating That the Four-Parameter Model Is Globally Identifiable

A sufficient requirement for global identifiability of a model is that the parameters be uniquely determined as a function of response-category probabilities (Erdfelder, in press). To show that this is the case for the four-parameter model, we now solve the remaining independent Equations A3, A4, A10, and A11 for each of the free parameters *P*, *M*, C_1, and C_2. Recall that the four-parameter model assumes that $M_1 = M_2 = M$.

The remaining independent equations, A3, A4, A10, and A11, are expressed in a simplified form in Equations A17, A14, A15, and A24, respectively. Substituting *M* for M_1 and M_2 in Equations A14 and A15 produces the following four independent equations:

$$p_{12} = (1 - C_1)(1 - c)(1 - p_{13}) \quad \text{(A17)}$$

$$p_{13} = PM + Pg - PMg \quad \text{(A26)}$$

$$p_{33} = Pg - PMg \quad \text{(A27)}$$

$$p_{41} = (1 - C_2)c(1 - p_{33}). \quad \text{(A24)}$$

The only independent equations containing the parameters *P* and *M* are the equations for p_{13} and p_{33}. We use these equations to solve for *P* and *M* in terms of the response-category probabilities. First, we subtract Equation A27 from Equation A26:

$$p_{13} - p_{33} = PM. \quad \text{(A28)}$$

Then, using Equation A28, we can make the following substitutions in Equation A26:

$$p_{13} = p_{13} - p_{33} + Pg - (p_{13} - p_{33})g. \quad \text{(A29)}$$

Equation A29 can be solved for *P*, resulting in the following equation expressing the parameter *P* as a function of response-category probabilities and the constant *g*:

$$P = p_{33}/g + (p_{13} - p_{33}). \quad \text{(A30)}$$

Equations A30 and A28 can be combined to solve for parameter *M:*

$$M = (p_{13} - p_{33})/P$$

$$M = (p_{13} - p_{33})/[p_{33}/g + (p_{13} - p_{33})]. \quad \text{(A31)}$$

By solving the independent equation that contains C_1, Equation A17, we obtain

$$C_1 = 1 - \{p_{12}/[(1 - c)(1 - p_{13})]\}. \quad \text{(A32)}$$

Similarly, Equation A24 can be used to solve for C_2:

$$p_{41} = (1 - C_2)c(1 - p_{33})$$

$$C_2 = 1 - \{p_{41}/[c(1 - p_{33})]\}. \quad \text{(A33)}$$

By substituting the appropriate constant values for *g* and *c* into Equations A30, A31, A32, and A33, we can express the parameters in terms of the response-category probabilities. Thus, the four-parameter model is globally identifiable.

Appendix B

Experiment 1: Response Frequencies as a Function of Instruction Condition and Item Type

	Response type		
Condition and item type	Y	N	PM
PMI			
Target, match	29	7	60
Target, nonmatch	4	28	64
Nontarget, match	720	165	11
Nontarget, nonmatch	61	824	11
CMI			
Target, match	45	7	44
Target, nonmatch	5	46	45
Nontarget, match	737	155	4
Nontarget, nonmatch	49	842	5

Note. $n = 32$ participants in each condition. Y = "yes" response; N = "no" response; PM = prospective memory response; PMI = prospective memory task important; CMI = color-matching task important.

Appendix C

Experiment 2: Response Frequencies as a Function of Instruction Condition and Item Type

Condition and item type	Response type Y	N	PM
PMI, same			
Target, match	6	3	66
Target, nonmatch	1	9	65
Nontarget, match	545	121	5
Nontarget, nonmatch	48	625	6
PMI, different			
Target, match	11	5	59
Target, nonmatch	1	16	58
Nontarget, match	564	107	2
Nontarget, nonmatch	33	643	1
CMI, same			
Target, match	14	1	60
Target, nonmatch	4	17	54
Nontarget, match	559	111	4
Nontarget, nonmatch	53	616	7
CMI, different			
Target, match	18	7	50
Target, nonmatch	1	24	50
Nontarget, match	587	86	1
Nontarget, nonmatch	52	623	1

Note. $n = 25$ participants in each condition. Y = "yes" response; N = "no" response; PM = prospective memory response; PMI = prospective memory task important; CMI = color-matching task important.

Appendix D

Experiment 3: Response Frequencies as a Function of Encoding Time and Item Type

Condition and item type	Response type Y	N	PM
Short encoding time			
Target, match	33	3	42
Target, nonmatch	5	25	48
Nontarget, match	613	102	13
Nontarget, nonmatch	48	667	13
Long encoding time			
Target, match	20	3	55
Target, nonmatch	2	18	58
Nontarget, match	631	92	5
Nontarget, nonmatch	72	653	3

Note. $n = 26$ participants in each encoding condition. Y = "yes" response; N = "no" response; PM = prospective memory response.

Appendix E

Experiment 4: Response Frequencies as a Function of Encoding Time and Item Type for the Three-Response Model

Condition and item type	Response type Y	N	PM
Short encoding time ($n = 25$)			
Target, match	29	4	42
Target, nonmatch	3	21	51
Nontarget, match	623	62	15
Nontarget, nonmatch	37	648	15
Long encoding time ($n = 26$)			
Target, match	16	2	60
Target, nonmatch	0	13	65
Nontarget, match	648	76	4
Nontarget, nonmatch	1	684	3

Note. The empty cell was replaced by 1 when modeling the data. Y = "yes" response; N = "no" response; PM = prospective memory response.

Appendix F

Experiment 4: Response Frequencies as a Function of Encoding Time and Item Type for the Four-Response Model

Condition and item type	Response type Y	N	Y, PM	N, PM
Short encoding time ($n = 23$)				
Target, match	25	4	34	6
Target, nonmatch	4	18	4	43
Nontarget, match	566	62	13	3
Nontarget, nonmatch	42	587	4	11
Long encoding time ($n = 22$)				
Target, match	13	2	46	5
Target, nonmatch	0	12	6	48
Nontarget, match	549	64	3	0
Nontarget, nonmatch	39	574	0	3

Note. Empty cells were replaced by 1 when modeling the data. Y = "yes" response; N = "no" response; Y, PM = "yes" response followed by prospective memory response; N, PM = "no" response followed by prospective memory response.

Received May 28, 2002
Revision received December 5, 2003
Accepted December 5, 2003 ■