Prospective Memory Decision Control: A Computational Model of Context Effects on Prospective Memory

Luke Strickland[1], Vanessa Bowden[2], & Shayne Loft[2]

[1]Curtin University

[2]The University of Western Australia

Author Note

Luke Strickland, Future of Work Institute, Curtin University; Vanessa Bowden, School of Psychological Science, The University of Western Australia; Shayne Loft, School of Psychological Science, The University of Western Australia.

This research was supported by DP200101842 from the Australian Research Council awarded to Loft and Strickland. Correspondence concerning this article should be addressed to Luke Strickland, Future of Work Institute, Curtin University, 78 Murray Street, Perth, Western Australia 6000, Australia (email: luke.strickland@curtin.edu.au).

We would like to thank Sarah Thorne for collecting the data. Study materials are available upon request. The experimental data and analysis code are available at: https://osf.io/qebmp/. This study was not preregistered.

Accepted 17/February/2023, *Journal of Experimental Psychology: Learning, Memory and Cognition*

Abstract

Prospective memory (PM) tasks require remembering to perform a deferred action and can be associated with predictable contexts. We present a theory and computational model, Prospective Memory Decision Control (PMDC), of the cognitive processes by which context supports PM. Under control conditions, participants completed lexical decisions. Under PM conditions, participants had the additional PM task of responding to letter strings containing certain syllables. Stimuli were presented in one of two colors, with color potentially changing after each set of four trials. A pre-trial colored fixation was presented before each set. Under control and PM standard conditions fixation color was meaningless. Under PM context conditions, fixation color indicated whether a PM target could occur within the next set. We replicated prior findings of higher PM accuracy for context compared to standard conditions, and the expected variation in PM costs (slowed lexical decisions) as a function of context relevance. PMDC, which formalizes PM as a process of evidence accumulation among ongoing and PM task responses, accounted for the impact of context on PM costs and accuracy via proactive and reactive cognitive control. Increased ongoing task thresholds and decreased PM thresholds in relevant contexts indicated proactive control. With context provision, PM accumulation rates on PM trials increased, as did inhibition of accumulation to competing responses, indicating reactive control. Although an observed capacity-sharing effect explained some portion of PM costs, we found no evidence that participants redirected more capacity from the ongoing to the PM task when contextually cued to relevant contexts.

Keywords: Prospective Memory, Context, Prospective Memory Decision Control, Computational Modeling

The completion of event-based prospective memory (PM) tasks requires remembering to perform a deferred action in response to an expected environmental event. PM is important for everyday life (e.g., remembering to give a friend a message when you next see them), can be critical for everyday safety (e.g., remembering to slow down when driving through a school zone; Bowden, Visser et al., 2017), and is essential in safety-critical work contexts such as aviation and healthcare (Dismukes, 2012; Loft et al., 2019).

Event-based PM is often examined in the laboratory using the Einstein and McDaniel (1990) paradigm, where participants perform an ongoing task (e.g., lexical decision), and are instructed to remember to make an alternative (PM) response when relatively infrequent target events are presented (e.g., a word containing the syllable 'tor'). This paradigm emulates that outside the laboratory individuals need to self-initiate the retrieval of a deferred action whilst engaged in other activities in the time between planning the action and when the action should be performed. Researchers have used variants of this paradigm to identify factors that promote prospective remembering and to develop theories regarding underlying cognitive processes (for review, see Rummel & McDaniel, 2019).

There are three prominent theories of PM. The Preparatory Attentional and Memory Processes (PAM) theory (Smith & Bayen, 2004) assumes the engagement of resource-demanding preparatory attention is required to be prepared to recognize a PM target, and thus that PM requires the allocation of cognitive capacity away from the ongoing task and toward a capacity-consuming PM target detection process (i.e., capacity sharing). Preparatory attention is indexed by higher ongoing task mean response times (RT) in blocks of trials in which individuals have a PM task relative to conditions when the ongoing task is performed alone (*PM costs*; Smith, 2003; also see Marsh et al., 2003). Positive correlations between PM accuracy and PM costs (Smith & Bayen, 2004), and findings such that emphasizing the importance of PM tasks increases costs and improves PM accuracy (Kliegel et al., 2004), are evidence used to support the notion that capacity sharing is functional to PM.

The Multi-Process View (MPV; Einstein et al., 2005, also see the Dynamic MPV proposed by Scullin et al., 2013) assumes that PM is often dependent on “monitoring” for PM targets, and thus that PM can require capacity sharing between the ongoing and PM task during stimulus processing. However, under some conditions, such as when PM targets are focal to ongoing tasks (i.e., a high overlap in information processing requirements), or when PM targets are highly associated with responses (e.g., say the word “pasta” if you see the word “spaghetti” during the ongoing task), the MPV assumes that attention to a PM target can lead to the reflexive (spontaneous) retrieval of the intended action, without capacity sharing and thus without PM costs (Einstein et al., 2005; McDaniel et al., 2004).

Researchers have also developed computational models to quantitatively account for outcomes from the Einstein and McDaniel (1990) paradigm (e.g., Boywitt & Rummel, 2012; Grünbaum et al., 2021; Heathcote et al., 2015b; for review see Strickland al., 2019a). Strickland et al. (2018) presented Prospective Memory Decision Control (PMDC), a theory and computational model able to identify the capacity sharing and cognitive control processes underlying both PM and ongoing task performance. PMDC is a linear ballistic accumulator (LBA) model (Brown & Heathcote, 2008) that formalizes PM as a process of evidence accumulation among independent ongoing task accumulators and a PM accumulator, in which the first accumulator to reach a threshold determines the response. PMDC instantiates the proactive and reactive cognitive control mechanisms specified by Braver’s (2012) dual-mechanisms theory. Contrary to the assumptions of PAM and MPV, the PMDC model indicates that PM costs in the Einstein and McDaniel (1990) paradigm are accounted for by proactive control over decision thresholds, and not capacity sharing (Strickland et al., 2019a).

Further, PMDC provides a comprehensive account of behavioral outcomes from two benchmark PM manipulations – target focality and PM importance – with proactive control of ongoing task and PM task response thresholds, and reactive control of evidence accumulation on PM trials, rather than variation in capacity sharing (Strickland et al., 2018).

The current study applies PMDC to provide a computational account of the cognitive processes by which contextual information can support PM. PM demands are often associated with predictable contexts. For example, an air traffic controller may hold the intention to assign any incoming aircraft type a non-routine alternate altitude and know that these aircraft typically approach the sector from a certain direction (Loft et al., 2011; Loft, 2014). In everyday life, we may hold the intention to pass a message to a work colleague, and have knowledge of when and where we are likely to encounter that colleague. Providing context to support participants' ability to predict when they will be able to perform the PM task in the Einstein & McDaniel (1990) paradigm can improve PM, and further, reduce PM costs in contexts not relevant to PM (see review in Bowden et al., 2021). PAM and the MPV account for context effects by assuming that individuals engage (increase) preparatory attention (PM monitoring) to detect PM targets in known relevant contexts, which is disengaged (decreased) in irrelevant contexts (Einstein et al., 2005; Scullin et al, 2013; Smith & Skinner, 2019). In the current paper we extend the literature by quantitatively testing, with PMDC, the relative roles of proactive control, reactive control and capacity sharing in accounting for the impact of context on PM costs and PM accuracy. Before describing the PMDC model in detail, we first review the literature on PM context.

Context Effects on Prospective Memory

The extent to which context improves PM is dependent on how context is manipulated (for further reviews of PM context effects see Bowden et al., 2021; Smith, 2017;

Smith & Skinner, 2019). When task instructions clearly demarcate the blocks of trials that can potentially contain a PM target event (PM relevant context) from those that cannot (PM irrelevant context), this typically results in (relative to PM conditions not provided context instructions) an elimination or reduction in PM costs in irrelevant blocks and/or increased PM cost in relevant blocks, and improved PM accuracy (e.g., Ball et al. 2015; Brewer & Marsh, 2010; Kominsky & Reese-Melancon, 2017; Meier et al. 2006; but see Ball & Bugg, 2018a).

In other studies, researchers have provided participants with progressive contextual information to allow them to anticipate the temporal proximity of relevant contexts (Bowden, Smith et al., 2017; Smith et al., 2017). Bowden, Smith et al. had participants make ongoing true/false judgements for sentences while remembering to respond differently to four studied words. PM-relevant contexts were cued by displaying progressive trial numbers. Participants who were told in advance the range of trial numbers in which PM targets could appear had reduced costs during irrelevant contexts, increased eye fixations to displayed trial numbers as relevant contexts approached, increased costs during relevant contexts, and improved PM accuracy. Smith et al. (2017) presented participants with campus photographs and asked them to remember to complete “errands” at specific locations. Presenting photographs in order, as though participants were walking around campus reduced costs and marginally improved PM accuracy relative to those presented photographs in random order because it allow participants to anticipate the temporal proximity of relevant contexts. In line with this, costs increased for the ordered condition as the PM-relevant target location neared.

Another approach to manipulating context associates the PM task with a trial type rather than with different blocks or ranges of ongoing task trials (e.g., Ball & Bugg, 2018b; Kuhlmann & Rummel, 2014; Lourenço & Maylor, 2014; Lourenco et al., 2013). For example, Lourenço et al. (2013) instructed participants that the PM target syllables would only be presented in either words or non-words (specific condition), or both (nonspecific condition). Although participants in the specific condition were aware that non-word letter strings were irrelevant to the PM task, it was not possible to predict before each trial onset whether the trial was relevant or not, and there was no improvement to PM accuracy.

Overall, a common element linking studies that have found a benefit to PM from context is that participants were provided with some degree of warning before presentation of relevant contexts. Bowden et al. (2021) systematically investigated how much warning was needed to benefit PM accuracy. Participants performed a lexical decision task while remembering to respond differently to PM targets (letter strings containing “tor”). Stimulus color was used to indicate “relevant” contexts in which a PM target could appear. Relevant context trials were presented in sets of four same colored trials, with the PM target appearing in either the first, second, or third trial position of relevant contexts. Bowden et al. found improved PM accuracy, including for PM targets presented in the first trial position of a relevant context, if relevant contexts were preceded by a 1-s pre-trial fixation (colored shape) that indicated whether an upcoming set of four trials was relevant to the PM task.

To account for the effects of contextual cuing observed in Bowden et al. (2021) and related studies, PAM and the MPV propose that capacity-consuming preparatory attention (PM monitoring) is engaged in relevant contexts. This increases PM accuracy, but also usurps capacity from the ongoing task, increasing PM costs. The theories agree that these processes could potentially be disengaged in irrelevant contexts, thereby decreasing PM costs. There is also agreement that warnings are required, or at least one trial spent in a relevant context, for context to improve PM because individuals need to identify PM relevant contexts in order to initiate preparatory attention (PM monitoring).

In addition to theorizing about the processes which are engaged in response to PM contexts, some theories also describe the processes that initiate the psychological response to contextual cues. Identifying PM contexts is argued to be a second event-based PM task (Kuhlmann & Rummel, 2014), where the context cue is the PM event and the response is a latent psychological shift that supports subsequent PM performance (e.g., initiating capacity-demanding monitoring for PM targets). There are several theoretical processes by which context cues might initially be identified. In the dynamic MPV (Scullin et al., 2013), some contextual cues can automatically trigger spontaneous retrieval of the PM intention (also see Lourenco & Maylor, 2014), and in turn trigger capacity-demanding monitoring. Similarly, in the PAM theory decisions about whether to engage capacity-consuming preparatory attentional processes are made at transition points between contexts (Smith et al., 2017). Guynn (2003)'s two-process theory specifies that participants holding PM intentions constantly maintain a capacity-consuming retrieval mode. This retrieval mode could potentially help identify PM contexts. Once contexts are identified, participants could then initiate target checks (Guynn, 2003). Ball and Bugg (2018a; 2018b) argued that context effects could be explained by dual mechanisms of attention control[1], with the type of attentional control mechanism and likelihood of context identification depending on the focality of the context cue and the overall task demand placed upon participants. Following context identification, Ball and Bugg (2018a; 2018b) assume that individuals allocate resource-demanding (capacity-sharing) monitoring processes.

[1] The control over attention specified by Ball & Bugg (2018a; 2018b) differs from the control over decisions specified by PMDC (Strickland et al., 2018). Both theories specify that there are proactive versus reactive modes of control, inspired by the distinction in Braver (2012). However, in PMDC, cognitive control slows ongoing task *decisions* relative to PM decisions to improve the likelihood of PM decisions. In contrast, the attentional control described by Ball & Bugg (2018a; 2018b) biases *attention* towards either PM or ongoing tasks, and thus in PMDC terms is more compatible with shifts in "capacity sharing".

Despite substantial theorizing about, and empirical investigation of, PM context effects, there is little associated computational modelling. The PMDC model (Strickland et al., 2018) can contribute in this regard. Although PMDC does not specify the mechanism by which context cues initiate a psychological response, it can identify the nature of the psychological response that follows PM context detection, with significant theoretical implications. Notably, existing verbal theories of PM context effects (Ball & Bugg, 2018a; 2018b; Einstein et al., 2005; Guynn 2003; Lourenco & Maylor, 2014; Scullin et al., 2013; Smith & Bayen, 2004) are uniform in contending that increases in PM costs and PM accuracy associated with PM-relevant contexts, and decreases in costs associated with irrelevant contexts, are due to strategic variations in capacity sharing between PM monitoring/preparatory processes and ongoing task processing. However, such effects could also be driven by cognitive control over decision making. PMDC can identify the contributions of each of these mechanisms to context effects through fitting to participants' choices and RT distributions. In the next section, we introduce the model.

Prospective Memory Decision Control

Most prior PM studies perform separate analyses of coarse manifest mean (average) measures of PM accuracy and PM cost. As a consequence, theories such as PAM, the MPV, and the DMPV have been developed without a formal account of the relationship between accuracy and RT, and with little reference to the skew or variability of RT distributions. This limitation is significant because the quantitative relationships between accuracy and RT distributions contains potentially important psychological information (Lerche & Voss, 2020). PMDC addresses these limitations by providing a unified account of the entire distribution of ongoing task RTs and choices, as well as the RT distributions and response choices observed on PM target trials. This is the major advantage of PMDC model analysis above other approaches.

As illustrated in Figure 1, PMDC assumes a race to response selection between PM and ongoing task decision processes. Each of the ongoing task accumulators and the PM accumulator has its own response threshold, which corresponds to the evidence that must be accumulated to make that decision. Upon the presentation of stimuli, evidence accumulates toward each decision at an accumulation rate, and the first to reach threshold determines the decision made (Brown & Heathcote, 2008). Thus, PM hits occur on PM trials where the PM accumulator reaches threshold before the ongoing task accumulators, whereas PM errors occur when the ongoing task accumulators reach threshold before the PM accumulator.

Fitting the PMDC model involves estimating the values of latent psychological quantities (i.e., parameters) that best account for the data, and examining how these latent quantities vary across experimental manipulations. PMDC is a process model with sound parameter recovery properties, and so its parameter estimates are informative for psychological theory (Heathcote et al., 2015a). Specifically, PMDC can recover the threshold and accumulation rate parameters key to psychological inference in feasible experiment designs similar to the current study (Strickland et al., 2018; 2021). Simulations demonstrate acceptably low estimation bias (i.e., small differences between Bayesian posterior mean estimates and "true" generating parameter values), and well-calibrated Bayesian uncertainty intervals [see, for example, p13. of Strickland et al. (2018)'s supplementary materials for plots of parameter recovery results]. Further, PMDC's parameters can be used to identify distinct underlying capacity sharing and cognitive control mechanisms [see p5-6 of Strickland et al. (2018)'s supplementary materials for a simulation demonstrating the model can distinguish capacity sharing from cognitive control with little mimicry].

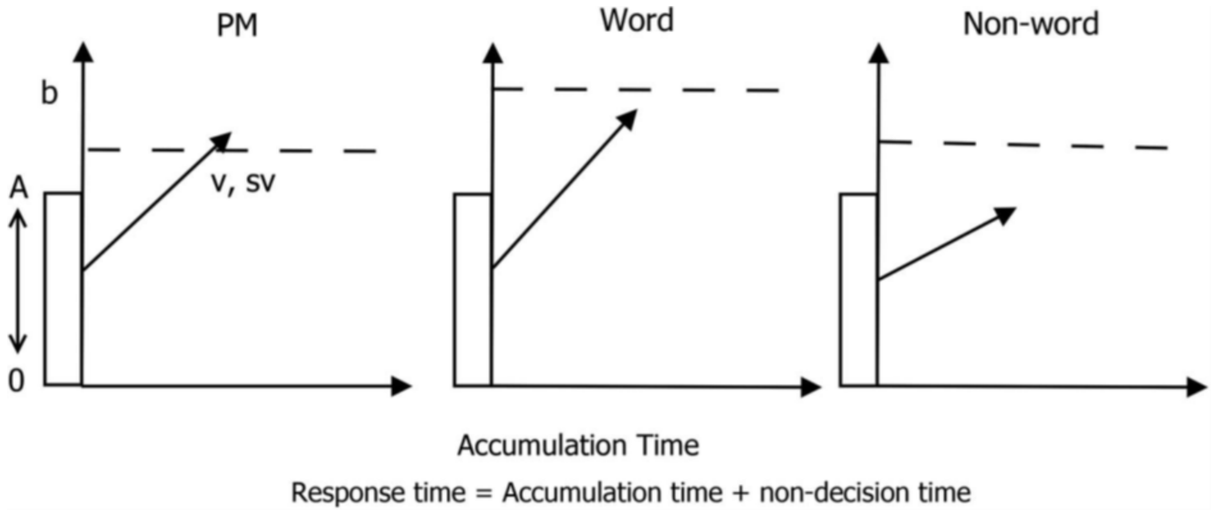

Figure 1. The PMDC Model (Strickland et al., 2018). Each accumulator begins with a starting amount of evidence drawn from the uniform distribution *U* [0, A]. Over time, evidence accumulates toward threshold *b* at a rate drawn from a normal distribution with mean *v*, standard deviation *sv*. The predicted response is determined by the first response accumulator to reach threshold. Total response time is equal to total time to accumulate to threshold, plus a non-decision time parameter included as time for additional processes such as stimulus encoding and motor responding.

PMDC measures capacity sharing between the PM task and the ongoing task processing by estimating the evidence accumulation rate parameters corresponding to non-PM trial responses. The idea, consistent with general theories of cognitive resources (e.g., Bundesen, 1990; Kahneman, 1973; Navon & Gopher, 1979; Pashler, 1984), is that processing speed of the ongoing task is increased in proportion to the amount of cognitive capacity it receives. Supporting this conceptualization, empirical work indicates that rates correlate with other measures of capacity (Donkin et al., 2014; Eidels et al., 2010) and are reduced with dual task load (Logan et al., 2014).

PMDC measures two types of ongoing task accumulation rates. *Match* accumulation rates reflect the accumulation toward the correct decision (e.g., the word accumulator's rate for a word in a lexical decision task), and *mismatch* accumulation rates reflect accumulation toward the incorrect decision (e.g., the non-word accumulator's rate for a word stimulus). For interpretability, it is helpful to convert the rates into theoretically oriented measures of processing capacity: processing *quality* (match accumulation rate – mismatch accumulation rate), and processing *quantity* (match accumulation rate + mismatch accumulation rate) (Boag et al., 2019a, 2019b). Processing quality indexes how accurate information processing is, with higher quality implying that evidence accumulation is better able to discriminate stimuli. In contrast, processing quantity indexes how much total evidence is accumulated per unit of time. With suitable experimental constraints, it is possible to test theories that make detailed predictions about attention mapping to processing quality and quantity (see Boag et al., 2019b). However, neither PAM theory, the MPV, or any other theory invoking PM monitoring specifies whether the PM task borrows resources that would affect ongoing task quality, or ongoing task quantity, or both. Thus, under these theories it is possible that PM costs could be associated with decreases in ongoing task processing quality and/or quantity.

Contrary to PAM and the MPV, evidence accumulation modelling often indicates that the PM costs are caused by increased thresholds, rather than decreased ongoing task capacity. In a standard paradigm (lexical decision), ongoing task accumulation rates generally do not vary substantially between control and PM conditions (e.g., Heathcote et al., 2015b; Horn & Bayen, 2015; Strickland et al., 2017, 2018). Further, elevations in costs under conditions that are argued to promote capacity sharing, such as when PM targets are non-focal or when PM importance is emphasized, have been found to result from shifts in thresholds rather than capacity (Strickland et al., 2018). PMDC does indicate that capacity sharing effects arise in more demanding tasks such as air traffic control (Boag, et al., 2019a, 2019b) and maritime surveillance (Strickland et al., 2019b). This is likely because capacity sharing occurs when capacity demands of a task approach an individual's capacity limit (Navon & Gopher, 1979), which is less likely to result from simpler ongoing tasks such as lexical decision.

Often, PM task responses must compete for response selection with routine ongoing task responses (Loft & Remington, 2010). The fact that PM responses are required less frequently than ongoing task responses puts PM responding at a disadvantage (Loft & Remington, 2013). PMDC assumes that in order to mitigate the habitual advantage of ongoing task responding over PM responding, the cognitive system invokes cognitive control over decision making. Cognitive control refers to the processes that allow humans to deviate from routine behavior and act in a goal-directed fashion (Miller & Cohen, 2001). In PMDC, cognitive control adaptively modifies the probability that ongoing task and PM accumulators win the race to threshold. PMDC can distinguish the effects of control from those of capacity sharing. Following related cognitive control concepts invoked in the PM literature (Bugg et al., 2013; McDaniel et al., 2013) and Braver's (2012) dual-mechanisms theory, PMDC includes both proactive control and reactive control.

Proactive control is activated in advance of an anticipated target event so that the control is applied when the event occurs to facilitate responding to that event (Braver, 2012). One form of proactive control is to increase ongoing task thresholds, so that when PM targets are presented ongoing task accumulators are less likely to reach threshold before and thus preempt the PM accumulator, increasing PM accuracy (Loft & Remington, 2013). Elevated ongoing task thresholds are a significant contributor to PM costs in every existing evidence accumulation analysis (for review see Strickland et al., 2019a). Numerous studies have also shown that elevated ongoing task thresholds can be the sole cause of PM costs (Heathcote et al., 2015b; Strickland et al., 2017, 2018), even when targets are non-focal or PM importance is emphasized (Strickland et al., 2018). However, simulations from PMDC (Strickland et al., 2018), and manipulations of ongoing task thresholds (Strickland et al., 2020), indicate that this form of proactive control provides minimal support to PM accuracy, as compared with

PMDC's other mechanisms. Proactive control can also be exerted over the PM response threshold. Decreasing the PM threshold increases the probability that the PM decision will reach threshold before the ongoing task decision. Strickland et al. (2018) reported lower PM thresholds when PM importance was emphasized or targets focal.

Reactive control processes occur at the time of a target event to facilitate the appropriate response to that event (Braver, 2012). Reactive control is activated by processing of PM stimulus features on PM trials (Figure 2). Specifically, PMDC theory proposes that processing of PM inputs activates a PM detector, which in turn triggers evidence accumulation toward the PM decision (reactive excitation; i.e., pathway A1 in Figure 2). Reactive excitation is estimated by the degree to which accumulation rates are faster on PM trials than non-PM trials. PM detector activation, and hence PM accumulation rates, could be affected both by stimulus-driven ("bottom-up") factors (the strength of PM stimulus inputs) and the attention/capacity allocated to processing PM inputs ("top-down" control). Consistent with reactive control driven by stimulus characteristics, Strickland et al. (2018) found greater PM accumulation rates for focal compared to non-focal targets, in line with the assumption of the MPV and neurological evidence (McDaniel et al., 2013) regarding greater bottom-up activation from focal PM stimuli. When stimulus characteristics are held constant, PM accumulation rates can be interpreted as a measure of the processing capacity devoted to PM. Indeed, Boag et al. (2019b) found evidence for such a process in a demanding environment which required capacity sharing between PM and ongoing tasks. Specifically, they found that PM accumulation rates were larger when PM task performance was emphasized, consistent with the allocation of greater cognitive capacity to the PM task.

PMDC also proposes that the PM detector has inhibitory connections to competing ongoing task decisions and thus reduces accumulation rates toward these decisions (reactive inhibition). This reactive inhibition is measured by the degree to which ongoing task accumulation rates are slower on PM trials than non-PM trials. Such inhibition could increase due to greater activation of the PM detector, with activation driven either by stronger PM-related stimulus features or biases in attention towards PM inputs. Consistent with this, Strickland et al. (2018) found increased inhibition of ongoing task responses was critical to driving the effects of PM focality. However, it is also possible for reactive inhibition to vary due to changes in control strategy, rather than changes in PM detector activation (and hence vary independently from reactive excitation). Specifically, participants may alter the sensitivity of inhibition to detector activation (weights associated with paths B1 and B2 in Figure 2). Strickland et al. found evidence that this mechanism drove the effect of PM importance on PM accuracy in a simple paradigm where there was little evidence of capacity sharing. Specifically, they found that when the importance of the PM task was emphasized, reactive inhibition increased without corresponding increases in excitation. Simulations suggested that this increased PM accuracy, as it prevented ongoing-task decisions from pre-empting the PM decision.

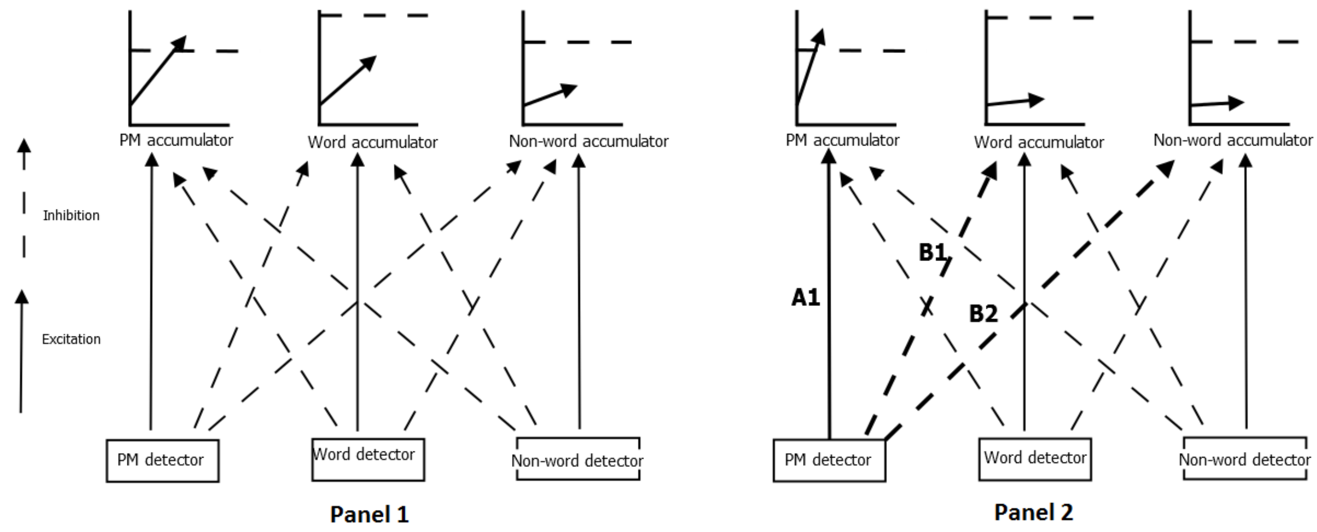

Figure 2. PMDC's Control Architecture (Strickland et al., 2018). Stimulus inputs activate the detectors, which in turn excite their corresponding accumulators, and inhibit other accumulators. Panel 1 depicts all the possible feedforward connections between detectors and accumulators. Processing PM inputs potentially activates the PM detector, which affects accumulation rates. Panel 2 shows how PM detector activation affects the decision process — an increase in the PM accumulation rate (via A1), and a decrease in ongoing task accumulation rates (via B1 and B2).

The Current Study: Prospective Memory Decision Control and Context

To investigate the effects of contextual cuing for PM tasks on the capacity sharing, proactive control and reactive control processes measured by PMDC, we combined the designs of Bowden et al. (2021) and Strickland et al. (2018). In control blocks, participants completed a lexical decision task. In PM standard blocks and PM context blocks[2], participants completed lexical decisions with the PM task of responding to letter strings containing a certain syllable. Stimuli were presented in one of two colors, with color potentially changing after each four-trial set. A 1s pre-trial contextually diagnostic fixation (colored rectangle) was presented before each four-trial set. In control and PM standard blocks, participants were instructed to ignore stimuli color as it was not relevant to either the

[2] Bowden et al. (2020) referred to the condition that received a warning of upcoming context and the "PM context + warning condition", and the condition that did receive a warning the "PM context condition". In the current study all conditions that received context received the warning, so for brevity we refer to this simply as "PM context blocks".

ongoing, and in the case of PM standard condition, the ongoing or PM task. In PM context blocks, participants were instructed that color indicated whether a PM target could occur on a particular set of four trials, and a PM target was presented on either the first, second, third or fourth trial position of some (but not all) relevant context sets.

We made several modifications to the Bowden et al. (2021) design, in line with Strickland et al. (2018) to enable reliable PMDC modelling. First, we used a within-subjects rather than between-subjects design because it improves the reliability of cognitive modelling. Second, Bowden et al. presented only a small number of PM targets and this would not produce enough data to reliably constrain a model of PM processes. We therefore increased the power of our model fitting by keeping the PM target trial to ongoing task trial ratio similar to Bowden et al. (approx. 1:19), but presenting more trials overall (1344 trials for each of three conditions, tested over three days). Third, as in Strickland et al. and other past studies applying PMDC, we instructed participants to rest their fingers on both ongoing task response keys (with one hand), and on the PM response key (with the other hand), so we could assume equal motor response time (Voss et al., 2010).

We expected to replicate the findings of Bowden et al. (2021). That is, we expected PM accuracy to be higher in PM context compared to PM standard blocks. Further, we expected PM costs to be lower in PM context compared to control/PM standard blocks during irrelevant sets, and higher for context compared to control/PM standard blocks during relevant sets. The novel contribution of our study is to then apply the PMDC model to distinguish between the contributions of capacity sharing and cognitive control to context effects on PM accuracy and PM cost.

PAM theory and the MPV both assume that PM costs are caused by, and PM performance supported by, capacity sharing between monitoring for PM targets and performing the ongoing task. They propose that contextual cueing stimulates more ongoing task capacity to be devoted to PM monitoring in relevant PM contexts, and less ongoing task capacity to be devoted to PM monitoring in irrelevant contexts. Thus, these theories would predict costs to ongoing task processing quality and quantity that are larger for relevant trial sets in context blocks than for matching trial sets in standard blocks, and smaller for irrelevant trial sets in context blocks than for matching sets in standard blocks. Strickland et al. (2018)’s PMDC model has previously indicated that PM costs to lexical decisions are not driven by shifts in ongoing task capacity, nor are variations in costs caused by focality and importance. Instead, proactive and reactive cognitive control over decision making drove these effects. As cognitive control allows humans to respond adaptively to their environment, it would follow that control processes are applied to adapt to changes in PM context relevance. Thus, it may be that the effects of context on PM accuracy and PM costs are driven by cognitive control, rather than (or in addition to) variation in capacity sharing.

To the extent that participants adjust proactive control based upon contextual cueing, they would be expected to decrease ongoing task response thresholds in PM irrelevant contexts relative to when in control/PM standard blocks (thereby decreasing PM costs), and increase ongoing task response thresholds in PM relevant contexts relative to standard blocks (increased PM costs). They would also be expected to lower their PM response thresholds compared to standard blocks in relevant contexts due to their contextual knowledge that PM targets can be presented. Participants might also apply more reactive control with appropriate contextual cueing, for example by adjusting the sensitivity of excitation and inhibition to PM inputs (i.e., weighting the connections depicted in Figure 2 more strongly). Increased excitation would be expected to increase PM accumulation on PM trials that were contextually cued relative to standard conditions. However, an increase in PM accumulation could also be explained by increased capacity devoted to the PM task, whether drawn from the ongoing task or from idle capacity. Another hallmark of control is reactive inhibition – that is, slower ongoing task accumulation on PM trials relative to non-PM trials. An increase in reactive control in context blocks would be expected to lead to further decreased ongoing task accumulation rates on PM trials (inhibition), compared to standard blocks.

Method

Participants. Forty-eight undergraduate students from The University of Western Australia participated in exchange for course credit. The study was approved by the University of Western Australia's Human Research Ethics Office.

Materials. The experiment was programmed in E-prime (Schneider et al., 2002). A pool of 1944 low-frequency (occurring 1-7 times per million) English words (length between 5 and 10 characters) were selected from the Sydney Morning Herald word data base (Dennis, 1995). A pool of 1944 nonword stimuli were created using the Wuggy algorithm (Keuleers & Brysbaert, 2010), which replaced subsyllabic segments of each word from the 1944 word pool with other subsyllabic segments (i.e., that are also legal in the same position). Wuggy replaces two of three subsyllabic segments and matched both the subsyllabic segment lengths and transition frequencies of its output nonwords with the input words. An additional 72 low-frequency (occurring 1-7 times per million) English words (length between 5 and 10 characters) served as PM word targets. Of these, 24 contained 'tor' (e.g., purgatory), 24 targets contained 'per' (e.g., groper), and 24 targets contained 'ver' (e.g., shiver). Finally, an additional 72 non-words (of length between 5 and 10 characters) from the Wuggy algorithm served as PM non-word targets. Of these, 24 contained 'tor' (e.g., pibtutory), 24 targets contained 'per' (e.g., criper), and 24 targets contained 'ver' (e.g., verinee).

Participants performed nine blocks of 448 trials – three control blocks, three PM standard blocks, and three PM context blocks – one block of each type per day over three days (total trials per condition = 1344). Condition (block type) was balanced across days such that participants did not receive a condition (e.g., PM standard) in the same position twice and such that each condition was presented equally often in each position over the three days (this resulted in 12 unique block orders counterbalanced over 48 participants). In control blocks, participants were presented with 224 words and 224 non-words. In PM standard and PM context blocks, participants were presented with 212 (non-target) words, 212 (non-target) non-words, 12 word targets and 12 non-word targets.

Over three days, each of the three target syllables (tor, ver, per) was used once in a PM standard block and once in a PM context block (no PM targets presented in control blocks). For example, on Day 1 a participant could have received ‘tor’ for a standard block and ‘ver’ for a context block, and on Day 2 ‘per’ for standard and ‘tor’ for context, and on Day 3 ‘ver’ for standard and ‘per’ for context.

Three pairs of stimuli font colors were created; red/green, yellow/purple, and blue/orange. The designation of color pairings to block type (control, PM standard, PM context) was counterbalanced but fixed for each participant over the entire three days.

For the context block, one of the colors in the assigned pairing was used for relevant sets and the other for irrelevant sets. Using the example of a red/green pairing, where red referred to the relevant context, participants were instructed that letter strings would be colored either red or green for four trials in succession (4-trial sets). They were further instructed they could reliably predict whether the upcoming 4-trial set would present strings of letters in red or green by the color of the fixation rectangle presented at the beginning of each 4-trial set. They were instructed that PM targets could be presented in red sets, but would not be presented in green sets. For the context block, there were 28 4-trial sets in which participants were told that targets could be presented (relevant sets). However, a target was only presented in 24 of these 28 relevant sets. Each block of 448 trials was precisely divided into 112 4-trial sets, and consisted of seven relevant sets (six that contained a target and one that did not) and 21 irrelevant sets. The assignment of relevant sets to the 112 trials was random with the exception that two relevant sets could not be presented consecutively. The 12 target words and 12 target non-words for each block were randomly assigned to trials in each relevant set that contained a target. Six targets were presented in position 1 (i.e., the first trial in a relevant set), six in position 2, six in position 3 and six in position 4.

For standard and control blocks, stimuli color had no relevance to PM target presentation. For each standard and control block, one designated color was presented on 75% of trials and the other color on 25% of trials.

Before control blocks, participants were instructed what to expect regarding the colored fixation and 4-trial sets, but were also told that the color of the letter strings was not important and to be ignored. No targets were presented in the control condition.

Before standard blocks, participants were instructed what to expect regarding colored fixation triangles and 4-trial sets, but were also told that the color of the letter strings was not important and to be ignored. They were also told that the color of the letter string had no relationship to whether a letter string would be a target. For standard blocks, six targets were presented in the designated color of the 4-trial sets that were presented 25% of the time, and 18 targets in the designated color of the 4-trial sets that were presented 75% of the time, thereby keeping the ratio of target presentation equal across the ratio of stimuli color presentation (i.e., 25%, 75%). This was implemented so that contexts (fixation and stimuli color) were uninformative in standard blocks, to avoid participants learning contextual associations over time.

Procedure. For the lexical decision task, participants were instructed that they would be presented with letter strings and to respond indicating whether strings were words or nonwords (e.g., press ‘s’ for word, ‘d’ for nonword) as quickly and accurately as possible. Depending on the upcoming block type, participants received different instructions. Before control blocks, participants were instructed that they only needed to make lexical decisions. Before standard and context blocks, participants were instructed to press an alternative response key instead of the word or non-word response key if presented a target word, for example, “In the next block of lexical decision trials, if you see a word or non-word with the syllable ‘tor’ in it then press ‘d’ instead of ‘j’’. We instructed participants to make their PM response instead of their ongoing task response (as done in previous PMDC studies) because it is relevant to everyday PM errors in which individuals perform a routine task action instead of a less common but required deferred task action (Norman, 1981; Reason, 1990). In addition, this means we record only one RT and one response on every trial, allowing us to fit PM RTs without confounds from the production of other responses.

Four response key assignments were counterbalanced across participants; (1) *s* = word, *d* = nonword, *j* = PM, (2) *d* = word, *s*= nonword, j =PM, (3) *k* =word, *j* = nonword, *d*= PM; (4) *j*= word, k= nonword, *d* = PM. The response key assignment remained the same for each participant across the entire experiment. Participants were instructed before each block (and after summative feedback, see below) to rest their fingers on designated response keys.

Each 4-trial set began with a colored fixation rectangle displayed on a black background for 1s. This was followed by a 250-ms blank screen before the first trial in the 4-trial set was presented. Each colored letter string remained on the black screen until a response key was pressed. If the participant made a correct word or nonword response (including a correct lexical decision response on a PM target trial, which is a PM miss), or a correct PM response, the next trial began after a delay of 1s with no other feedback. If the participant made an incorrect response the word ‘INCORRECT’ was presented in white for 1s, after which the subsequent trial began. After the fourth trial in each set, the next colored fixation rectangle was presented and the above procedure repeated.

Each day, participants first completed 24 practice lexical decision trials and received summative feedback on their accuracy (e.g., “90% correct”). A fixed set of additional 36 low frequency words and 36 non-words were selected from the Sydney Morning Herald word data base/Wuggy algorithm, and 12 words and 12 non-words used from this pool were used on each day for practice. Participants then proceeded to the experimental blocks and were presented with control, standard, or context instructions. Participants next completed a 3-min distractor puzzle, after which they began the first block of experimental trials. In addition to the trial-by-trial feedback for incorrect responses, after completion of each quarter of a block (i.e., 112 trials), participants were presented summative feedback on task accuracy. To avoid cueing the PM intention, a correct ongoing task response and a correct PM response counted toward the ‘correct’ status of trials in summative feedback. In addition to the breaks within blocks for summative feedback, participants were instructed to rest for 2 min between blocks.

Study materials are available upon request. The experimental data and analysis code are available at: https://osf.io/qebmp/ (Strickland, 2023). This study was not preregistered.

Results

We first report conventional analysis of ongoing task and PM performance to assess whether our manipulations had the expected effects. We examined the effects of block (control, PM standard, PM context), context (relevant trials, irrelevant trials), and stimulus type (word, non-word). In addition, to capture practice and fatigue effects, we included a "session" factor (session 1, session 2, session 3).

One participant's data was excluded due to fast responding (over 20% of responses <0.2s in several blocks), and another due to a block of low lexical decision accuracy (<60%). A third participant's data was excluded because of an error resulting in missing data for one of their blocks. The first trial of each block, and after each break, were excluded from subsequent analysis, as was any trial immediately following a PM trial and/or PM response. Trials with outlying RTs, defined as < 0.2s, or > mean + 3*IQR/1.349 (IQR/1.349 is a robust equivalent to standard deviation), were also excluded (4.64% of trials). Because of their rarity, we did not independently analyze "false alarms" (PM responses on non-PM trials; 0.1% of responses), or incorrect responses on PM trials (1.67% of PM trials, 0.06% of all trials). PM false alarm trials were excluded from analysis of ongoing task accuracy.

To analyze PM and ongoing task performance (accuracy and mean RT), we applied mixed-effects models including a random intercepts term for participants, and no random slopes. Accuracies were analyzed with a generalized linear model with a probit link function, and mean correct RTs were analyzed with a linear mixed model. We tested each experimental factor for significance using type II Wald's chi-square tests. The results of these tests are tabulated in the supplementary materials. Where appropriate, we followed up with planned contrasts using the emmeans package (Lenth et al., 2018), with Bonferroni adjustments for multiple comparisons. For all frequentist tests, we applied a conservative alpha of $p < .005$ (Benjamin et al., 2018). Standard errors reported in text and graphs were calculated using the Morey (2008) bias corrected method.

Ongoing Task

Accuracy. Ongoing task accuracies are displayed in Figure 3. There was an effect of stimulus type, with responses more accurate to non-words than words. This effect interacted with the session factor. Accuracy on word trials decreased from session 1 (*M* = 0.923, *SE* = 0.011) to session 2 (*M* = 0.914, *SE* = 0.011) and session 3 (*M* = 0.909, *SE* = 0.011), but accuracy on non-word trials did not (session 1 *M* = 0.946, *SE* = 0.011; session 2 *M* = 0.946, *SE* = 0.011; session 3 *M* = 0.945, *SE* = 0.011).

There was an interaction between stimulus type, block, and context. In control blocks, there were not significant differences in accuracy across relevant (non-word *M* = 0.942, *SE* = 0.008; word *M* = 0.912, *SE* = 0.009) and irrelevant trials (non-word *M* = 0.945, *SE* = 0.008; word *M* = 0.917, *SE* = 0.009). Similarly, in standard blocks there were not significant differences in accuracy across relevant (non-word *M* = 0.944, *SE* = 0.008; word *M* = 0.915, *SE* = 0.008) and irrelevant trials (non-word *M* = 0.946, *SE* = 0.008; word *M* = 0.917, *SE* = 0.008). In contrast, in context blocks word accuracy was higher for relevant trials (*M* = 0.931, *SE* = 0.009) than for irrelevant trials (*M* = 0.911, *SE* = 0.008). There was also a trend (p~.03) indicating non-word accuracy was lower for relevant trials (*M* = 0.938, *SE* = 0.009) than for irrelevant trials (*M* = 0.949, *SE* = 0.008). However, we are not aware of a strong reason why contextual cuing would be expected to increase word accuracy or decrease non-word accuracy in a PM task including both word and non-word PM targets, and this was a small effect, so it should perhaps be viewed with some skepticism.

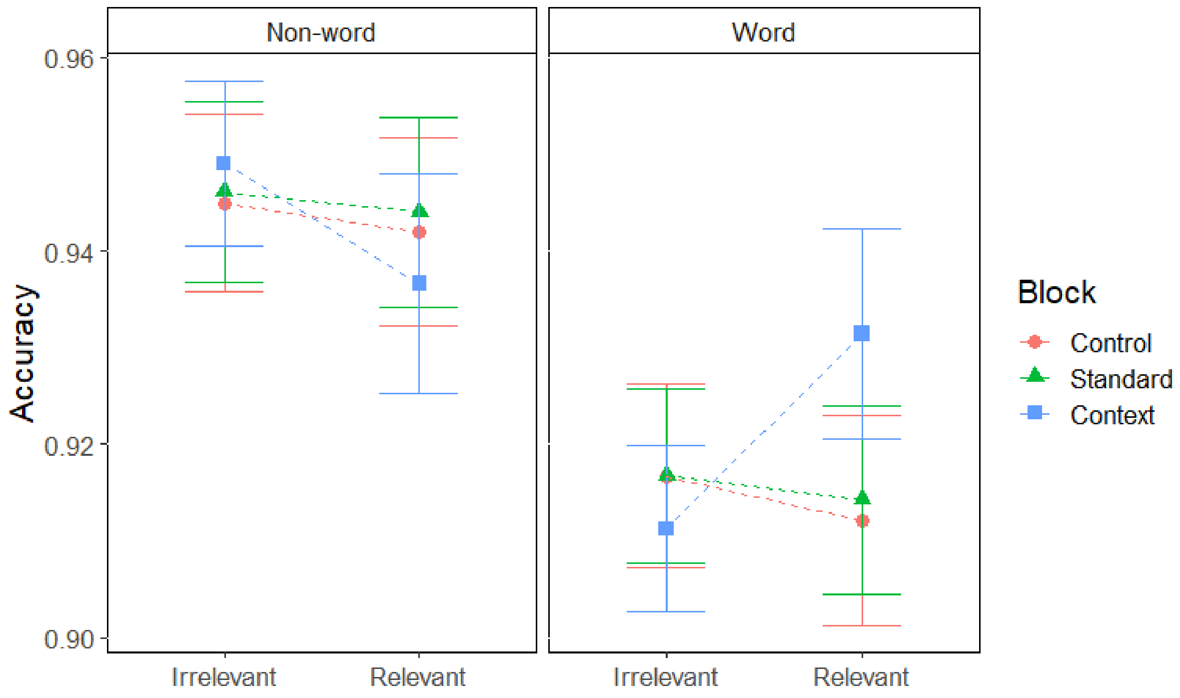

Figure 3. Ongoing task accuracies. Each panel corresponds to one stimulus type. The error bars included were calculated using the Morey (2008) bias-corrected method for within-subjects error bars.

Response Time. Correct ongoing task RTs are displayed in Figure 4. There was an effect of stimulus type, with responses slower to non-words (*M* = 0.808s, *SE* = 0.018s) than words (*M* = 0.750s, *SE* = 0.016s). There was a main effect of condition, with RTs slower in standard blocks (*M* = 0.819s, *SE* = 0.005s) than control blocks (*M* = 0.743s, *SE* = 0.004s), and intermediate in context blocks (*M* = 0.783s, *SE* = 0.004s). This effect interacted with trial context. RTs were not significantly different between relevant and irrelevant contexts in the control (relevant *M* = 0.747s, *SE* = 0.004s; irrelevant *M* = 0.742s, *SE* = 0.005s) or standard (relevant *M* = 0.820s, *SE* = 0.005s; irrelevant *M* = 0.819s, *SE* = 0.005s) blocks. In contrast, in context blocks relevant trial RTs (*M* = 0.867s, *SE* = 0.005s) were much slower than irrelevant trial RTs (*M* = 0.767s, *SE* = 0.005s). Due to this interaction, RTs were slower to relevant trials in context blocks than standard blocks, and RTs were faster to irrelevant trials in context blocks than standard blocks. There was still some cost in context blocks on irrelevant trials as compared with control blocks, however this effect was relatively small. There was also an interaction between block and session, with PM costs decreasing on later sessions, but remaining significant even on session three (see supplementary materials).

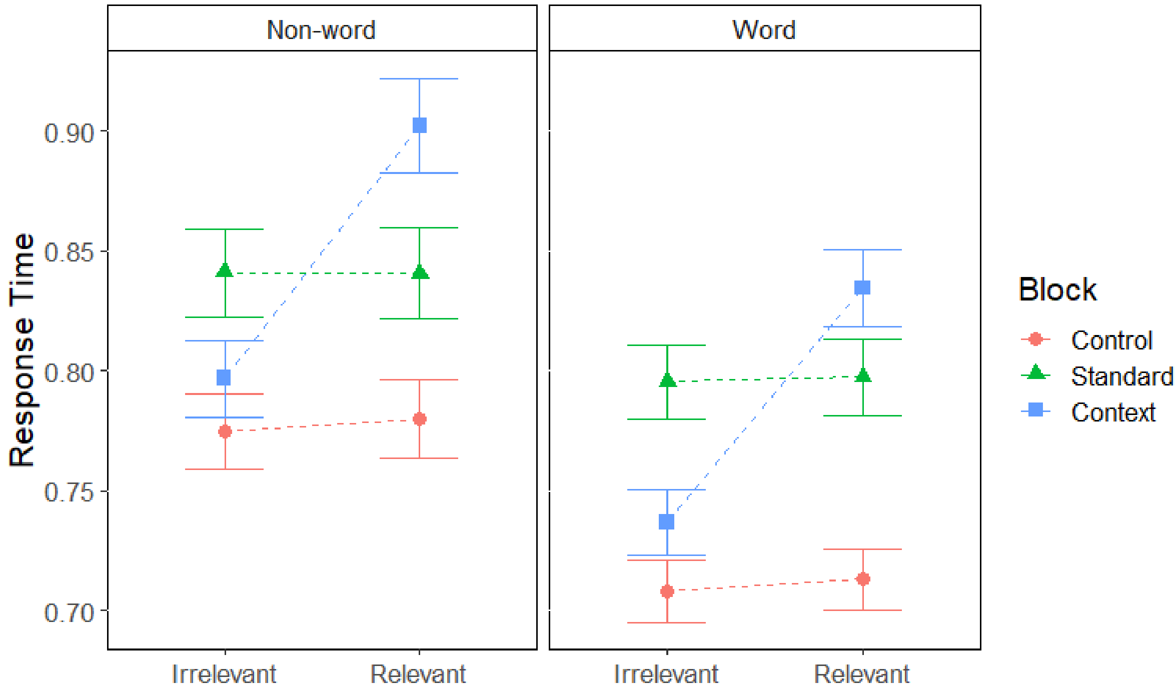

Figure 4. Mean correct ongoing task response times. Each panel corresponds to one stimulus type. The error bars included were calculated using the Morey (2008) bias-corrected method for within-subjects error bars.

PM Task

Accuracy. PM task accuracies are displayed in Figure 5 as a function of block and stimulus type. There was a main effect of block, with PM accuracies higher in context than standard blocks. The only other effect reaching significance was a main effect of session. PM accuracy was significantly lower on session 1 (M = 0.731, SE = 0.036) than on session 2 (M =

0.772, *SE* = 0.032), but neither was significantly different to session 3 (*M* = 0.757, *SE* = 0.032).

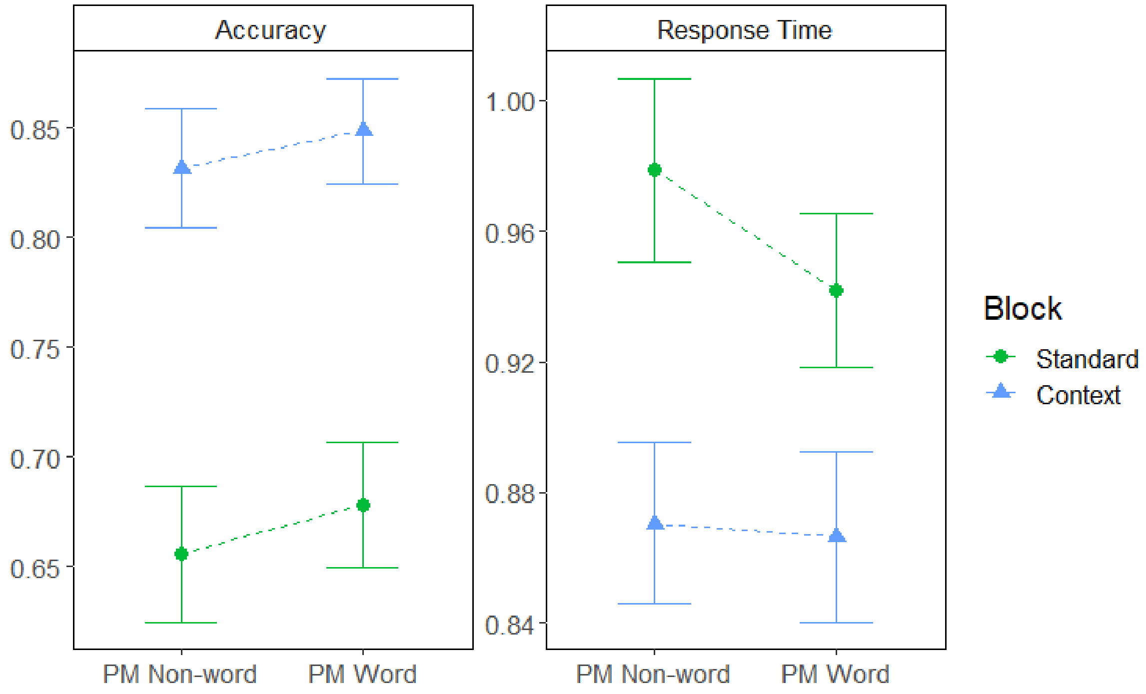

Figure 5. PM task accuracies (left panel) and mean correct response times in seconds (right panel). The error bars included were calculated using the Morey (2008) bias-corrected method for within-subjects error bars.

Response Time. Correct PM RTs are displayed in Figure 5 as a function of block and stimulus type. There was a main effect of block, with PM RTs faster in context blocks than standard blocks. The only other effect that reached significance was a main effect of session. PM RTs were significantly slower on session 1 (*M* = 0.952s, *SE* = 0.032s) than on session 2 (*M* = 0.890s, *SE* = 0.025s) and Day 3 (*M* = 0.900s, *SE* = 0.027s), with session 2 and session 3 not significantly different to each other.

To summarize, we replicated Bowden et al. (2021). PM costs were reduced for context blocks compared to standard blocks in irrelevant sets and were slower for context blocks than standard blocks in relevant sets. PM accuracy was higher, and correct PM responses faster, in context compared to standard blocks.

Model Results

The core underlying model was a three-accumulator LBA (Figure 1). Model parameters could potentially vary for each accumulator (word, non-word, PM), block (control, standard PM, context PM), context (relevant trial, irrelevant trial), stimulus type (word, non-word, PM word, PM non-word), and session (session 1, session 2, session 3). However, we implemented a range of plausible parameter restrictions prior to fitting, for theoretical reasons and to accommodate good parameter estimation.

We estimated only one start-point noise parameter and one non-decision time parameter[3], and accumulation rates were drawn from truncated normal distributions bounded above 0, both consistent with previous applications of PMDC (e.g., Strickland et al., 2018). We estimated only one "*sv*" parameter indexing rate variability for "matching" rates (including PM rates on PM trials), across all conditions, and the "mismatching" parameter was fixed at 1. These parameter constraints were implemented because they improve parameter estimation properties and computational speed, and yet allow adequate flexibility to test the theoretical questions of interest. As we report later, our observed fits suggest that we fit a flexible enough model to describe the observed data.

We estimated thresholds in terms of B (which is $b - A$). Thresholds were allowed to vary for different accumulators, blocks, contexts, and experiment days. They were not allowed to vary for different stimulus types: as participants were not aware what stimulus would be presented, it would be circular to assume they modify thresholds based on stimulus

[3] In response to a reviewer comment, we fitted a model with different non-decision times for each condition (control/standard/context), for each context (relevant trials/irrelevant), to check for evidence of an additional PM target check which extends non-decision time (Horn & Bayen, 2015). For example, Guynn (2003)'s theory specifies that PM target "checking" may temporarily occur either before or after making an ongoing task decision. As reported in the supplementary materials, this model was not supported by model comparison. Further, the estimated non-decision times were not consistent with target checking. Thus, we focus in text on the simpler model with one non-decision time per participant. Alternative parameter inferences from the more-flexible model are available in the supplementary materials.

type. Further, thresholds for the PM accumulator were not estimated separately for relevant and irrelevant trials. In context blocks (where thresholds would be expected to vary), there were no PM trials in irrelevant contexts, and in PM standard blocks the designation of ‘irrelevant’ sets versus “relevant” sets had no functional meaning (i.e., the stimulus color differed but all trials were all “relevant” to PM).

For non-PM trials, we estimated separate mean accumulation rates for each ongoing task accumulator, stimulus type, block and by context trial type (relevant/irrelevant). Because we observed very few PM “false alarms” (responses on non-PM trials), we estimated a single accumulation rate for these responses for each participant. For PM trials, we estimated a separate evidence accumulation rate for the PM accumulator for each stimulus type and block. Rather than estimate ongoing task accumulation rates for PM trials, the differences between ongoing task accumulation rates on relevant non-PM trials and on PM trials were determined by “inhibition” parameters, which we estimated. To determine the mean ongoing task accumulation rate on PM trials, the inhibition parameter was subtracted from the appropriate non-PM trial accumulation rate for “relevant” trial sets. For example, accumulation towards the “word” decision on a PM word trial would be determined by subtracting the “word” inhibition parameter from the accumulation rate towards “word” on (relevant) non-PM word trials. Inhibition for the “matching” accumulator was estimated separately for each condition and each stimulus type. Because “mismatching” lexical responses on PM trials (e.g., “non-word” response on a PM word trial) were rarely observed, inhibition for the mismatching accumulator was poorly estimated. Thus, we estimated only a single inhibition parameter for mismatching lexical responses across blocks and stimulus types for each participant.

We estimated model parameters using Bayesian methods from the Dynamic Models of Choice suite (Heathcote et al., 2019). Specifically, we applied a form of Markov chain Monte Carlo, which samples parameter values proportionality to the posterior probability of those parameter values. The details of estimation, which are generally similar to Strickland et al. (2018), are available in the supplementary materials.

Model Fit

Figures 6 and 7 display the posterior predictions of the model. Overall, the model provided a close fit to the observed trends in accuracy and RT, including the effects of the context manipulation.

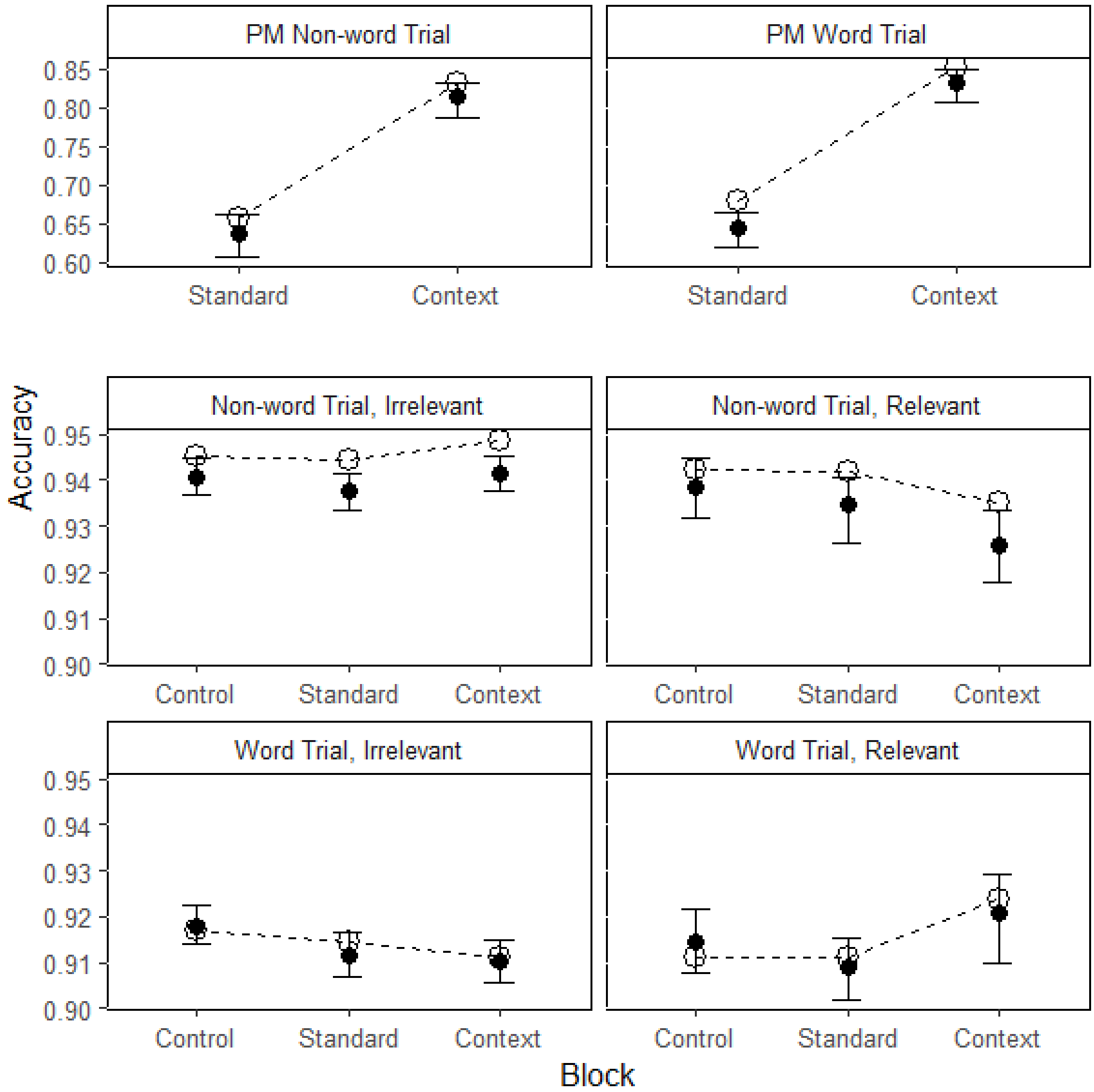

Figure 6. Posterior predictions for accuracies, averaged over participants. The model predictions correspond to the white circles, the posterior means correspond to the black shaded dots. The error bars display the 95% posterior credible intervals of the predictions.

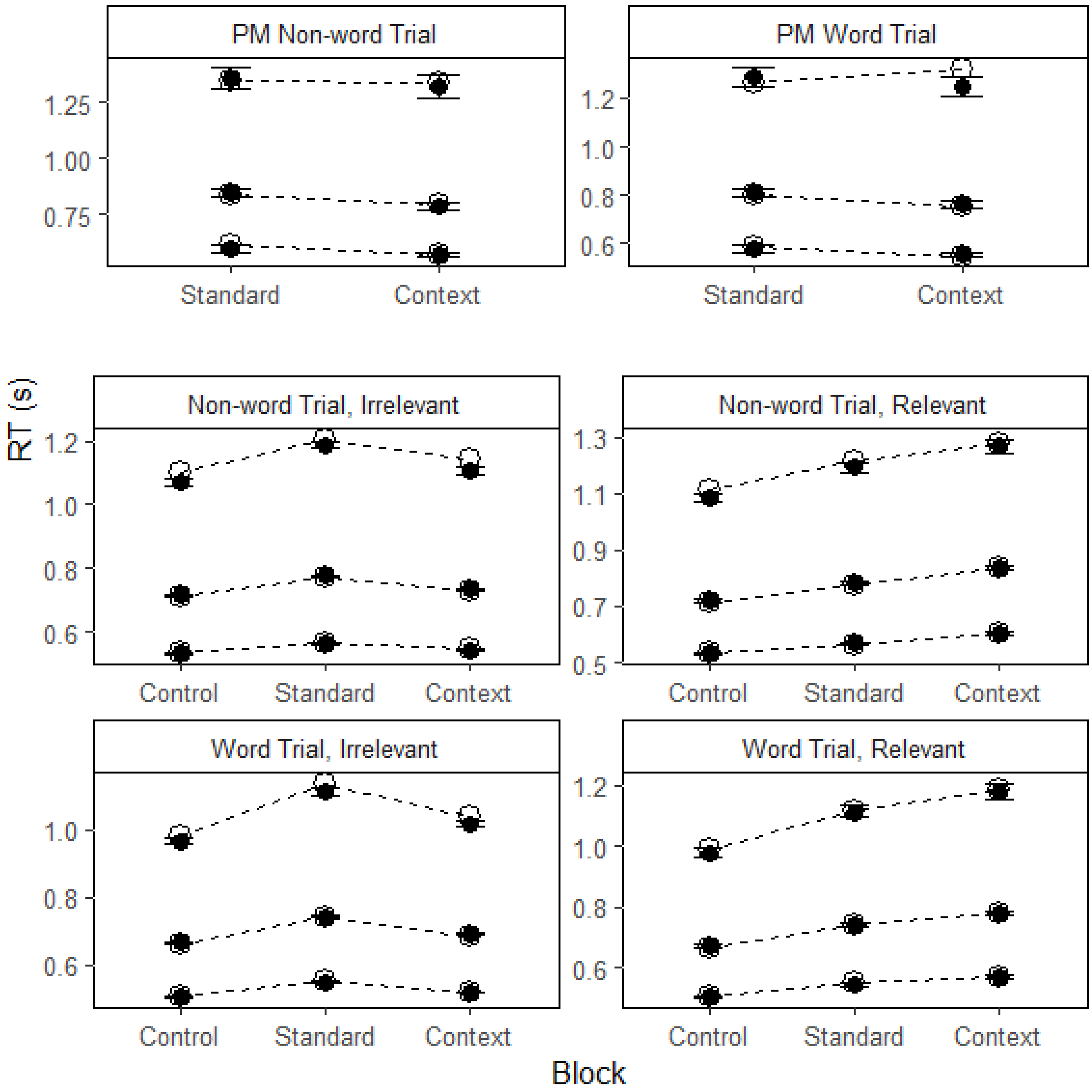

Figure 7. Posterior predictions for response time (RT), pooled over participants, and over correct and error responses. The model predictions correspond to the white circles, the posterior means correspond to the black shaded dots. The error bars display the 95% posterior credible intervals of the predictions. Three quantiles of RT are depicted. The bottom quantiles on each plot represent the 0.1 quantile, the middle the median RT, and the top the 0.9 quantile of RT.

Model Mechanisms

To examine parameter estimates, we created a 'subject-averaged' distribution, which averaged the values of each parameter over all subjects for each posterior sample. The posterior mean of the *A* parameter was 0.18 (*SD* = 0.01), the posterior mean of the non-decision time parameter was 0.15 (*SD* = 0.0017), and the posterior mean of the *sv* parameter towards decisions matching the correct response was 0.52 (*SD* = 0.004).

In subsequent sections, we review how parameters vary across blocks, focusing on the effects of context. To statistically test parameter differences, we use a posterior *p* value based on the number of times that one parameter was sampled at a higher value than the other. We report the posterior *p* in the direction against observed effects, to be consistent with common intuition about *p* values. Thus, if we observed parameter x was mostly larger than parameter y, we report posterior *p* as the proportion of samples on which y was larger than x. Likely due to our large trial numbers, many posterior *p*s were <.001. For an estimate of effect size we examined a posterior *Z* score, which was the posterior mean of the effect divided by its standard deviation.

We used simulations to explore how model mechanisms related to our model's predictions of PM costs and PM accuracy effects. To do so, we systematically removed model mechanisms and examined resulting miss-fit. To remove model mechanisms, we set parameter values from one of our experimental blocks equal to another block (e.g., setting PM standard block parameter values equal to control blocks). To the extent that removing a parameter difference causes miss-fit to an effect, that parameter difference drove the model's prediction of that effect. We report our simulations in detail in the supplementary materials. For brevity, in text we only report the percentage of the observed effect that is predicted (posterior mean) when model mechanisms are removed. To evaluate the extent to which predictions were diminished with the model mechanism removed, these percentages are compared with the percentage predicted by the "full model" with no mechanisms removed. Lower percentages compared with the full model indicates that the removed mechanism was more important in driving the effect, whereas higher percentages indicate the removed mechanism did not drive the effect. The reported percentages can be >100%, in the event that removing the mechanism makes the effect larger than was observed, or negative in the event that removing the mechanism makes the model predict an effect with the opposite sign.

Capacity Sharing

The estimates of non-PM trial accumulation rates are depicted in Figure 8. To test capacity sharing, we examined processing quality (the difference between matching and mismatching accumulation rates) and processing quantity (the sum of matching and mismatching accumulation rates). For interpretability, we report results aggregated over stimulus type in text, but comparisons split by stimulus type are tabulated in the supplementary materials.

The tests of capacity sharing are in Table 1. Comparing standard and context PM blocks with control blocks, we found some evidence of capacity sharing: processing quantity was lower than control, for both relevant and irrelevant trials. However, inconsistent with capacity sharing, there was evidence of improved processing quality in both standard and context blocks compared to control. Thus, the net effect of PM blocks on ongoing task capacity was not clear, although larger effect sizes for the quantity effect compared to quality effect point towards an overall reduction in ongoing task capacity with PM demands.

Our simulations assessing the extent to which capacity sharing drove our model's predictions of PM costs suggested that capacity effects did contribute, but to a lesser degree than proactive control. The full model predicted 98% of standard block costs to relevant trials, 97% of standard block costs to irrelevant trials, 99% of context block costs to relevant trials, and 88% of the very small context block costs to irrelevant trials. Removing capacity sharing diminished predicted costs, with the model predicting 67% of standard costs to relevant trials 73% of standard costs to irrelevant trials, 86% of context block costs to relevant trials, and 48% of the (small) context block costs to irrelevant trials.

We found that evidence for differences in capacity sharing between standard and context blocks was inconsistent. Crucially, on relevant trial sets, there were not notable differences between ongoing task accumulation quality or quantity in context blocks compared with standard blocks. Consistent with this, simulations indicated that capacity sharing effects did not play a role in our model's ability to predict increased PM costs on relevant trial sets for context blocks. Removing differences in capacity sharing actually led to a larger effect predicted than observed (121%), contrasting with the full model which predicted the effect accurately (101%). Thus, in context blocks participants did not redirect more ongoing task capacity towards the PM task when they were contextually cued on relevant trial sets, compared to standard blocks. However, on irrelevant trial sets, there was one relatively small effect (compared with the effect sizes of other mechanisms) in line with less capacity sharing for irrelevant trial sets in context blocks than in standard blocks (higher quantity in context blocks).

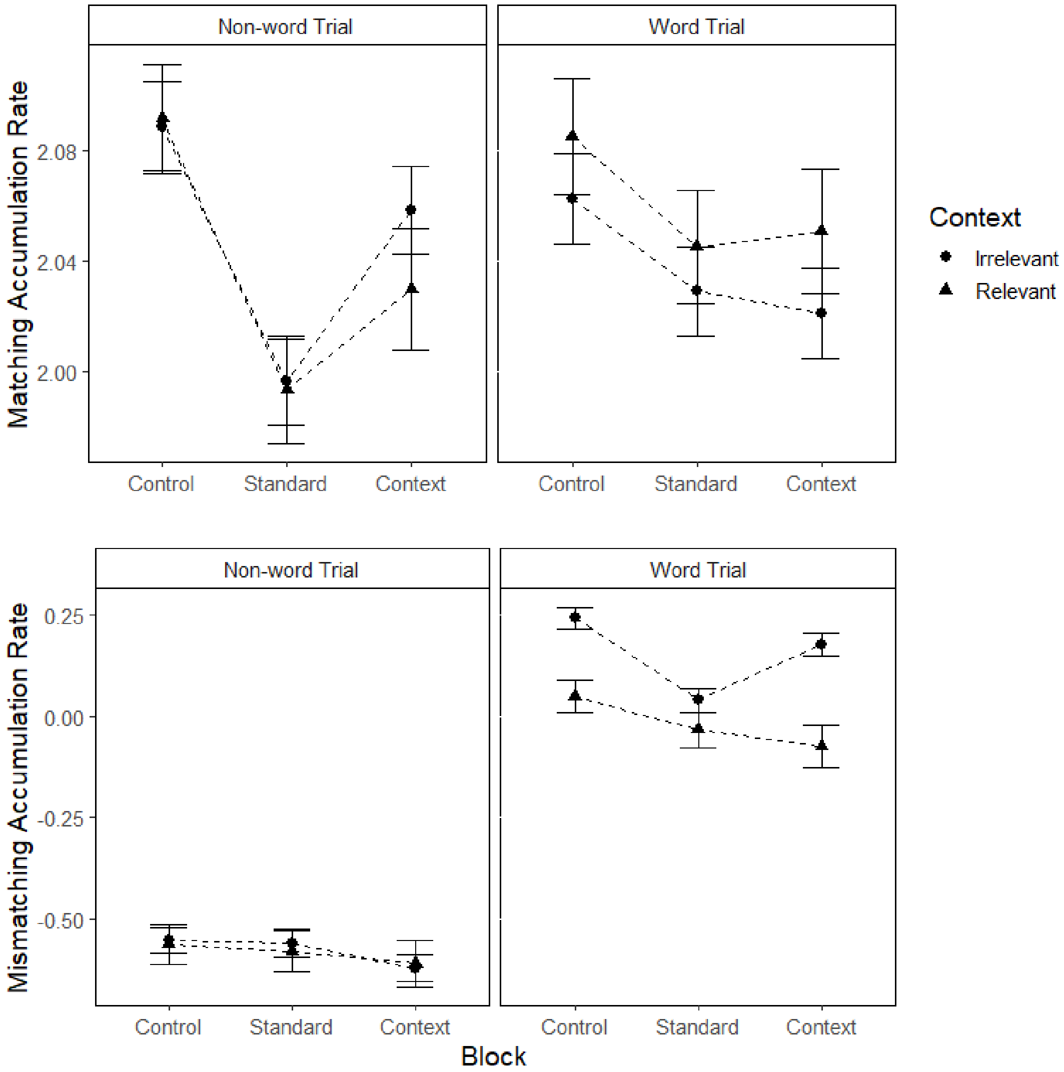

Figure 8. Estimates of accumulation rates. The shapes indicate the posterior means and the error bars correspond to the mean plus or minus the posterior standard deviation.

Table 1. *Statistical tests of capacity sharing on non-PM trials for both standard and context conditions. We report effects in terms of Z (p), where Z refers to the mean of the posterior samples of the effect divided by its standard deviation, and p refers to the proportion of posterior samples below 0 (for positive effects) or above 0 (for negative effects). Quantity refers to the sum of ongoing accumulation rates, and quality to the difference between the match and mismatch accumulator.*

Comparison	Quality	Quantity
Control - Standard (Irrelevant)	-2.08 (0.02)	5.97 (<.001)
Control - Standard (Relevant)	0.52 (0.299)	2.31 (0.011)
Control - Context (Irrelevant)	-1.62 (0.052)	3.82 (<.001)
Control - Context (Relevant)	-0.87 (0.191)	2.35 (0.009)
Standard - Context (Irrelevant)	0.52 (0.302)	-2.29 (0.011)
Standard - Context (Relevant)	-1.30 (0.096)	0.26 (0.398)

Proactive Control

Threshold estimates are depicted in Figure 9, and tests of proactive control in Table 2. There was evidence of proactive control over ongoing task thresholds in the standard and context blocks, with thresholds larger than in control blocks. Notably, there was a large effect of context on ongoing task thresholds in context blocks. Ongoing task thresholds were much higher in relevant contexts in context compared to standard blocks. In contrast, thresholds were lower in irrelevant contexts in context compared to standard blocks, although still higher than control block thresholds in irrelevant contexts.

Simulations indicated that proactive control over ongoing task thresholds were the most significant contributor to our model's predictions of PM costs, and essentially the only contributor to the additional PM costs associated with relevant trial sets in context blocks. With proactive control removed from the model, it predicted only 29% of standard costs to relevant trials, 21% of standard costs to irrelevant trials, 14% of context block costs to relevant trials, and 32% of the context block costs to irrelevant trials. Further, the model's predictions of increased costs on relevant trial sets in context conditions as compared with standard conditions were completely undermined (the effect predicted was -6% of the observed, as compared with 101% in the full model). We also found that removing proactive control over ongoing task thresholds diminished predictions of the PM accuracy advantage in the context condition. A model with differences in proactive control over ongoing task thresholds removed predicted 82% of the observed PM accuracy effect, as compared to the full model that predicted 105% of the observed.

There was also evidence of proactive control over the PM threshold: it was lower in context blocks, where participants were forewarned they were in a relevant context, than in standard blocks, $Z = 6.80$, $p < .001$. Simulations indicated that this mechanism contributed moderately to our model's predictions of higher PM accuracy in context blocks than standard (removed PM threshold effects led to predicting 82% of the observed effected, as compared with 105% in the full model). Further, it contributed substantially to model predictions of decreased PM RT in context blocks compared with standard blocks (removed PM threshold effects led to predicting 20% of the observed effected, as compared with the full model which predicted an effect 94% of the observed).

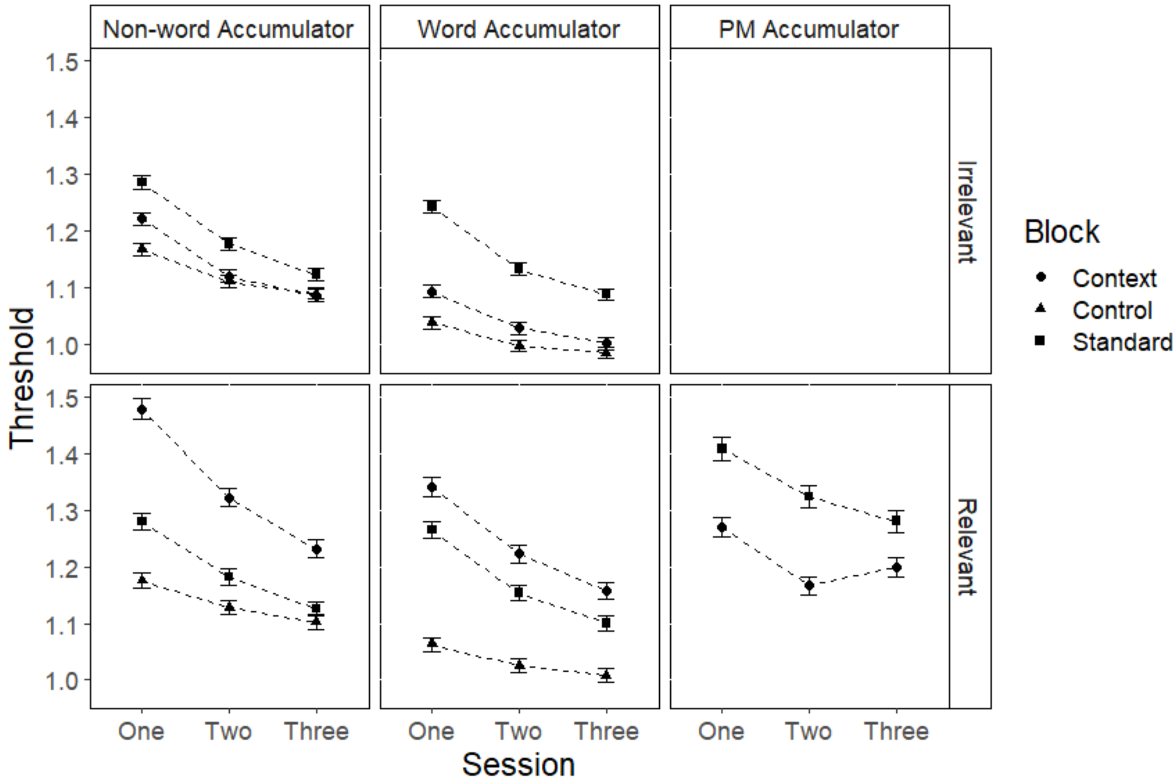

Figure 9. Estimates of thresholds. The shapes indicate the posterior means and the error bars correspond to the mean plus or minus the posterior standard deviation.

Table 2. *Statistical tests of proactive control over ongoing task decisions, which corresponds to the threshold in PM conditions minus its value in control conditions. We report effects in terms of Z (p), where Z refers to the mean of the posterior samples of the effect divided by its standard deviation, and p refers to the proportion of posterior samples below 0 (for positive effects) or above 0 (for negative effects). The combined column gives the size of the effect when samples of word and non-word thresholds are averaged prior to calculating the effect, which tended to increase certainty in the effects. This column is included for comparability with the analyses of ongoing task capacity, which are pooled across stimulus type.*

Comparison	Word Accumulator	Non-word Accumulator	Combined
Control - Standard (Irrelevant)	21.34 (<.001)	10.42 (<.001)	22.35 (<.001)
Control - Standard (Relevant)	11.76 (<.001)	5.22 (<.001)	12.14 (<.001)
Control - Context (Irrelevant)	5.46 (<.001)	3.05 (0.001)	6.00 (<.001)
Control - Context (Relevant)	15.14 (<.001)	14.65 (<.001)	21.35 (<.001)
Standard - Context (Irrelevant)	16.46 (<.001)	7.47 (<.001)	16.75 (<.001)
Standard - Context (Relevant)	-4.61 (<.001)	-9.98 (<.001)	-10.45 (<.001)

Reactive Control

Unsurprisingly, PM accumulation rates were much larger on PM trials than non-PM trials, where they were very low consistent with low 'false alarm' rates (posterior M = -3.02, SD = 0.07). For word PM targets, PM accumulation rates were higher in context blocks (posterior M = 1.93, SD = 0.02) than standard blocks (posterior M = 1.81, SD = 0.03), Z = 3.86, p < .001. Similarly, for non-word targets, PM accumulation rates were higher in context blocks (posterior M = 1.83, SD = 0.02) than standard blocks (posterior M = 1.66, SD = 0.03), Z = 5.03, p < .001. Simulations demonstrated that these effects contributed moderately to the model's predicted advantage of PM accuracy in context blocks, with a model removing excitation effects predicting 83% of the observed effect (compared with the full model's prediction of 105%). Further, differences in excitation contributed quite substantially to the model's predicted speeding of PM RT, a model with excitation removed predicted only 33% of the effect, down from 94% in the full model.

We found evidence of reactive inhibition, consistent with Strickland et al. (2018). With the model parameterization we applied, the inhibition parameter tracks the difference between ongoing task accumulation rates on non-PM trials, as compared with PM trials, and thus positive values indicate lower ongoing task accumulation rates on PM trials than non-PM trials. Inhibition to ongoing task word accumulation was substantial (much > 0) on PM word trials in context blocks (posterior M = 1.69, SD = 0.05), and in standard blocks (posterior M = 1.21, SD = 0.03), and there was stronger inhibition in context compared to standard blocks, Z = 8.37, p < .001. Inhibition to ongoing task non-word accumulation was substantial on PM non-word trials in context blocks (posterior M = 1.64, SD = 0.05), and in standard blocks (posterior M = 1.25, SD = 0.03), and there was stronger inhibition in context compared to standard blocks, Z = 6.34, p < .001. Inhibition for 'incorrect' lexical decision responses on PM trials, which was estimated combining all PM trials due to low observation numbers, was also substantial (posterior M = 1.47, SD = 0.07).

Simulations indicated that increased inhibition was a relatively strong contributor to the PM accuracy advantage in context blocks. Removing inhibition led to the model predicting only 70% of the observed PM accuracy effect, down from 105%. However, the simulations also indicated that increased inhibition detracted from the observed increased correct PM RT in context blocks (that is, without the inhibition effects, the speedup in correct PM RTs would have been even more substantial with contextual cueing). Specifically, a model with inhibition removed predicted an RT effect 129% of that observed, compared with 94% in the full model. This is consistent with statistical facilitation (if ongoing task accumulators are faster on average, they are more likely to outrun the PM accumulator).

Model Comparison

In summary, PMDC parameter estimates indicated that both capacity sharing and proactive control contributed to PM costs, but proactive control effects were significantly larger, and only proactive control contributed substantially to the effects of context on PM costs. Contextual cuing was associated with multiple shifts in parameters supporting PM accuracy: proactive control over ongoing task thresholds, proactive control over PM thresholds, faster accumulation rates towards the PM decision (excitation), and stronger inhibition of ongoing task accumulation on PM trials.

We previously discussed simulations in which we eliminated model mechanisms by setting parameter values equal across experimental blocks after parameter estimation to examine how those mechanisms contributed to our model's predictions of PM costs and PM accuracy. A complimentary approach is to estimate (fit) simpler models that disallow model mechanisms from the outset and examine whether those models provide a favorable tradeoff between parsimony and fit. Restricting model mechanisms during parameter estimation allows other mechanisms (e.g., proactive control) to account for data (e.g., PM costs) that would have been claimed by the excluded mechanism in a more flexible model.

We examined a series of restricted models using the deviance information criterion (DIC; Spiegelhalter et al., 2012). DIC selects models based upon a tradeoff between parsimony and fit, with a lower DIC suggesting that a model would achieve higher out-of-sample predictive accuracy. There are strong arguments against relying heavily on statistical information criteria when building cognitive theories (Navarro, 2019), and DIC values do not necessarily select the “true” generating model. However, DIC does provide information about differences in the overall predictive accuracy of competing models (McElreath, 2020). The magnitude of differences in DIC is meaningful. They can be viewed in terms of model weights, with a difference over 10 considered substantial (corresponding to a >99% preference for the model with the lower DIC).

We assess differences in DIC to get a picture of the importance of model mechanisms to fit relative to their costs to parsimony. We systematically excluded PMDC model mechanisms and compared the restricted models to the ‘top’ flexible model reported in the sections above (summed DIC = -56510). We report DICs in terms of differences from the ‘top’ model [i.e., DIC (top model) – DIC (restricted model)], with a lower (e.g., more negative) value indicating more support for the top model, and hence a higher importance of the excluded mechanism to predictive accuracy. DIC was calculated for each individual participant, and then summed.

We first examined the mechanisms underlying PM costs. To test whether excluding capacity sharing from our model could be justified, we fitted a model that estimated one set of non-PM trial accumulation rates across control, standard and context blocks. We found that such a model performed worse than the top (fully flexible) model (DIC difference = -36). We also examined a model that did not allow proactive control over ongoing task thresholds, by fixing ongoing task thresholds over control, standard and PM blocks. This model was far less supported than the top model (DIC difference = -5294), and also far less supported than the model that excluded capacity sharing. These results imply that whilst including capacity sharing effects was supported by model comparison, including proactive control over ongoing task thresholds was far more important for predictive accuracy.

We also examined the influence of model mechanisms that primarily affected PM trials – proactive control over PM thresholds, and reactive control over accumulation rates. Because PM trials were relatively rare, these mechanisms were expected to have smaller effects on DIC. A model excluding differences in proactive control over PM thresholds across conditions, by estimating one PM threshold that applied across context and standard conditions, had a substantially less favorable DIC than the top model (DIC difference = -381). Similarly, a model that excluded differences in reactive inhibition across standard and context conditions was disadvantaged relative to the top model (DIC difference = -71), although the relative advantage provided by retaining differences in inhibition was smaller than the advantage associated with differences in PM thresholds.

Interestingly, a model that excluded differences in PM accumulation rates (excitation) across conditions was favored by DIC above the top model (DIC difference = 156). Consequently, we examined effects in the DIC-preferred model (supplementary materials). We found that tests of all other model mechanisms yielded similar results to the top model, except that decreases in PM thresholds associated with context conditions were more substantial in the DIC-preferred model. This is likely because with differences in PM accumulation rates excluded, PM thresholds would need to decrease further in context conditions to fit the observed increased PM accuracy and deceased PM RT.

Discussion

PM demands are often linked to predictable environmental context. We applied the PMDC model to computationally account for the cognitive processes by which context supports PM. We adapted the Bowden et al. (2021) variant of the Einstein and McDaniel (1990) paradigm to enable accurate measurement of psychological processes. We replicated the Bowden et al. findings of improved PM accuracy and faster PM RT in context compared to standard blocks, of reduced PM costs in context compared to standard blocks during irrelevant sets, and of increased PM costs in context compared to standard blocks during relevant sets. These outcomes are what would be predicted by PAM theory, the MPV, and other theories of PM (e.g., Ball & Bugg 2018a; 2018b; Guynn, 2003) based on their assumption that individuals increase preparatory attention (monitoring) to detect PM targets in identified relevant contexts, and decrease monitoring in irrelevant contexts (Scullin et al, 2013; Smith & Skinner, 2019). However, analysis with the PMDC model indicated that contrary to these accounts, the effects of context were driven largely by variations in proactive and reactive cognitive control over decision making, rather than capacity sharing.

Capacity Sharing and Context

We found that ongoing task capacity costs were smaller in irrelevant contexts for the context condition, as compared with the standard condition, as would be predicted by PAM theory and the MPV. However, our model-based results were not consistent with capacity-sharing theories of context effects. We did not find that participants redirected more ongoing task capacity to PM monitoring when contextually cued to relevant PM contexts in context blocks as compared with standard blocks. Instead, increased PM costs in relevant contexts were driven by shifts in proactive control (increased ongoing task thresholds). As PM targets only occurred in relevant contexts in the context condition, this implies capacity sharing did not underlie the benefits to PM accuracy associated with context cues. We did find higher processing quantity for irrelevant contexts in context blocks, compared with standard blocks, which could be consistent with lower capacity sharing. However, this effect was small relative to associated proactive control effects, and because it occurred on irrelevant trials it is not directly relevant to the PM accuracy advantage of the context condition.

Although we found relatively little evidence that capacity sharing varied as a function of contextual cueing, we found some inconsistent evidence of overall capacity sharing. Processing quantity was lower for both PM blocks than control (in line with capacity sharing) in both relevant and irrelevant trial sets, but for irrelevant trial sets, processing quality was actually higher in both PM blocks than control (opposite direction implied by capacity sharing). Whilst inconsistent, this is stronger evidence of capacity sharing than has been found in our previous analysis of simple laboratory PM paradigms (e.g., Heathcote et al., 2015b; Strickland et al., 2017, 2018). However, it is worth noting that these capacity-sharing effects were smaller than threshold effects, less favored by formal model comparison, and accounted for a smaller proportion of PM costs. Further, we did not observe shifts in capacity sharing with contextual cuing of relevant PM contexts, and hence it could not have played a role in the effect of context on PM accuracy. Still, on balance the observed effects do provide some evidence that it is possible for capacity sharing to contribute to PM costs in a simple lexical decision paradigm. Anderson et al. (2018) also reported costs to ongoing task (lexical decision) accumulation rates, but as per our findings, a substantial portion of costs that they observed were attributed to shifts in ongoing task thresholds rather than decreased capacity.

Cognitive Control and Context

PMDC assumes that to mitigate the habitual advantage of ongoing task responding over PM task responding, the cognitive system invokes cognitive control (Braver, 2012). Proactive control was exerted over the ongoing task response thresholds, with increased thresholds in PM conditions relative to control. Proactive control effects were large, consistent, accounted for the majority of PM costs, and strongly supported by model comparison with DIC (Heathcote et al., 2015b; Strickland et al., 2018). Critically, proactive control exclusively accounted for the additional costs on relevant trials in context compared to standard blocks, and had a role in explaining the benefit of context to PM accuracy. This indicates proactive control of ongoing task thresholds does not only reflect an increase in caution due an increased overall perception of task complexity (Horn & Bayen, 2015), but rather that it can improve PM because it allows the PM accumulator more time to reach response threshold before the more routine ongoing task response (Loft & Remington, 2013; Heathcote et al., 2015b; but see conflicting findings by Strickland et al., 2020). Decreasing PM task thresholds serves the same function by increasing the probability that the PM decision will reach threshold before the ongoing task decision. We found that individuals did lower their PM thresholds in known relevant contexts during context blocks as compared with standard blocks. This effect was strongly supported by model comparison, and played a role in predicting the benefit of context to PM accuracy and PM RT.

In terms of PMDC, increased PM accumulation rates are consistent with increased PM excitation (activation of PM accumulation by PM evidence). Unsurprisingly, PM accumulation rates were larger on PM trials than non-PM trials. Further, estimates of PM accumulation rates were larger in context compared to standard blocks, and simulations suggested that this could underlie a portion of the increases in PM accuracy and faster PM

RTs observed with contextual cuing. However, in contrast to our other findings, this effect was rejected by formal model comparison, suggesting that a simpler model excluding differences in PM accumulation rates across conditions would provide better predictive accuracy (i.e., a better trade-off between parsimony and fit). We believe this indicates that the estimated differences in PM accumulation rates should be viewed with some skepticism.

Although our findings regarding increased PM accumulation (excitation) in context conditions were ambiguous (favored by estimation but not DIC), there are plausible psychological reasons it could occur. Contextual information was presented in the color of the stimuli, and so PM stimuli presented in PM relevant colors could provide stronger input to the decision process (color + syllable rather than just syllable). This idea is similar to the assumption of the Dynamic MPV that spontaneous retrieval can be prompted by the onset of PM relevant context, which in our case would be the PM relevant colored fixation (Scullin et al., 2013; Shelton & Scullin, 2013; also see related theories on context identification by Ball & Bugg, 2018a; 2018b; Smith et al., 2017). One difference is that these theories assume that participants, after context identification, subsequently increase “monitoring” for PM targets, decreasing capacity for the ongoing task and hence increasing costs. Our findings are at odds with this claim because they indicate that capacity sharing was not responsible for the increase in costs in relevant contexts in context blocks. Alternatively, increases in monitoring could potentially rely on reserve capacity – capacity not devoted to either the ongoing or PM tasks in other blocks (Strickland et al., 2019b; Boag et al., 2019a; Boag et al., 2019b).

We found reactive inhibition effects, consistent with Strickland et al. (2018). The “inhibition” parameters we estimated, which quantified the difference in ongoing task accumulation rates between PM trials and non-PM trials, were strongly positive. Further, they were larger in context than standard blocks. This difference was supported by model comparison, and removing inhibition led to a substantial decrease in the effect of contextual cuing on PM accuracy. This finding, that inhibition can be adjusted with appropriate contextual information, is consistent with the assertion that it is a form of cognitive control. Specifically, participants may have increased the strength of their inhibitory response to PM inputs with contextual cuing, as they have done when PM tasks are important (Strickland et al., 2018). Notably, the observed effects of inhibition are not consistent with the original specific assumption of the MPV or Dynamic MPV that spontaneous retrieval alone can prompt the PM response, without the need to inhibit the ongoing task decision (but see Bugg et al., 2012 for an integration of verbal cognitive control theory with MPV).

An alternative explanation for decreased ongoing task accumulation on PM trials we observed (measured via inhibition parameters) is that participants engaged in “acute” capacity sharing, that occurred only on PM trials. This would differ from the capacity sharing specified in PAM, MPV and other theories of PM in that it is not required for monitoring in advance of PM targets. Instead, processing PM target features could lead participants to allocate subsequent attention at the decision stage away from ongoing task accumulation and towards the PM task. However, this explanation predicts greater PM accumulation rates in context conditions, which was not supported by model comparison. Further, the explanation is inconsistent with findings from Strickland et al. (2018). They found that ongoing task accumulation rates decreased more on PM trials in a (supposedly less capacity demanding) focal PM task than on PM trials in a non-focal task. They also found that ongoing task accumulation rates were lower on PM trials in an important PM task than an unimportant PM task, but without any accompanying increase in PM accumulation (excitation). Thus, attributing the current decreased ongoing task accumulation rates PM trials to reactive inhibition is most consistent with our findings and with previous literature.

Conclusions and Future Research

In summary, the PMDC model accurately described both PM and ongoing task performance in a paradigm that included contextual cuing of PM events. It did so with increased ongoing task thresholds and decreased PM thresholds in PM relevant contexts, consistent with participants applying proactive control to adapt contextual information. In addition, inhibition of accumulation to competing responses increased with contextual cuing, consistent with enhanced reactive control of decision making. Although we found that an overall capacity-sharing effect explained some portion of PM costs (as compared to control blocks), there was no evidence that capacity sharing drove the effects of context on PM (differences between context and standard PM blocks). Thus, our findings pose challenges to a wide range of theories of context effects, such as the PAM theory and the MPV, which propose that context effects are driven primarily by variation in capacity sharing between monitoring for PM targets and performing the ongoing task.

Although our model speaks in detail to the latent processes engaged in response to identified PM context cues (e.g., proactive control over decision making), it is less informative regarding the processes responsible for initiating the psychological response to context. These processes are described by verbal theories of PM such as the dynamic MPV (Scullin et al., 2013), PAM theory (Smith, 2017), and dual mechanisms of attentional control theory (Ball & Bugg, 2018a, 2018b), with a focus on extent to which the processes underlying context identification are capacity demanding. Given that context cues were very obvious in our task (color of a large pre-stimulus fixation rectangle) and presented at a time uncoupled with other cognitive demands (before stimulus presentation), identifying contextual cues was likely not particularly capacity demanding. An important future direction is to specify a computational model of the processes that initiate the psychological response to contextual cues. For example, a context monitoring process could be specified that controls whether PMDC mechanisms (e.g., proactive control) respond to context cues. Because context detection is not directly observed but inferred by its effect latent psychological processes (e.g., proactive control), identifying the processes responsible will likely be difficult. Experimental manipulations targeting context effects, such as context cue focality and task demands, may provide constraints that help identify such a model. As suggested by Ball and Bugg (2018a), neuroscientific approaches are likely to play a crucial role in understanding the latent processes underlying context identification and how they compare to the retrieval processes that are activated when PM events are presented. Such data could potentially be integrated with PMDC using emerging techniques from computational cognitive neuroscience (Forstmann et al., 2016).

People are not always explicitly informed of the association between relevant contexts and PM tasks as they were in our paradigm, and yet it could still be adaptive to learn the relevance of such contextual information. Thus, a direction for future research will be to investigate how individuals learn associations between PM contexts and targets over time. To date, learning processes underlying PM have received little attention, with the exception of Strickland et al. (2022), who applied PMDC to model the learning of multiple targets over the target repetitions the effects on PM decision processes. Strickland et al. indicated that PMDC can be used as a meeting point between in quality of knowledge representations (the strength of connections between presented cues and knowledge representations such as context; Nosofsky & Palmeri, 1997; Ratcliff et al., 2004), and the dynamic PM decision process. We would expect that even without context instructions, individuals could learn to associate PM tasks with specific contexts that they expect to encounter (Kuhlmann & Rummel, 2014).

Such processes could potentially be explained by models combining PMDC with formal instantiations of reinforcement learning (Sewell et al., 2019; Miletić et al., 2020).

References

Anderson, F. T., Rummel, J., & McDaniel, M. A. (2018). Proceeding with care for successful prospective memory: Do we delay ongoing responding or actively monitor for cues? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*, 1036–1050.

Ball, B. H., Brewer, G. A., Loft, S., & Bowden, V. (2015). Uncovering continuous and transient monitoring profiles in event-based prospective memory. *Psychonomic Bulletin & Review, 22*, 492-499.

Ball, B. H., & Bugg, J. M. (2018a). Context cue focality influences strategic prospective memory monitoring. *Psychonomic Bulletin & Review, 25*, 1405-1415.

Ball, B. H., & Bugg, J. M. (2018b). Aging and the strategic use of context to control prospective memory monitoring. *Psychology and Aging, 33*, 527-544.

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., ... & Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour, 2*, 6-10.

Boag, R., Strickland, L., Heathcote A., Neal, A., & Loft, S. (2019a). Cognitive control and capacity for prospective memory in complex dynamic environments. *Journal of Experimental Psychology: General, 148*, 2181-2206.

Boag, R. Strickland, L., Loft, S, & Heathcote, A. (2019b). Strategic attention and decision control support prospective memory in a complex dual-task environment. *Cognition, 191*, 1-24. [Article 103974].

Bowden, V. K., Smith, R. E., & Loft, S. (2017). Eye movements provide insights into the conscious use of context in prospective memory. *Consciousness and Cognition, 52*, 68-74.

Bowden, V., Smith, R. E., & Loft, S. (2021). Improving prospective memory with contextual cueing. *Memory and Cognition, 49*, 692-711.

Bowden, V. K., Visser, T. A. W., & Loft, S. (2017). Forgetting induced speeding: Can prospective memory failure account for drivers exceeding the speed limit?" *Journal of Experimental Psychology: Applied, 23*, 180-190.

Boywitt, C., & Rummel, J. (2012). A diffusion model analysis of task interference effects in prospective memory. *Memory & Cognition, 40*, 70–82.

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences,* 16, 106–113

Brewer, G. A., & Marsh, R. L. (2010). On the role of episodic future simulation in encoding of prospective memories. *Cognitive Neuroscience, 1*, 81-88.

Brown, S., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57*, 153–178.

Bugg, J. M., McDaniel, M. A., & Einstein, G. O. (2013). Event-based prospective remembering: An integration of prospective memory and cognitive control theories. In D. Reisberg (Ed.), *The oxford handbook of cognitive psychology* (pp. 267–283). Oxford, England: Oxford University Press

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523–547

Dennis, S. (1995). The Sydney Morning Herald word database. Noetica: Open forum. Retrieved from http://psy.uq.edu.au/CogPsych/Noetica

Dismukes, R. (2012). Prospective memory in workplace and everyday situations. *Current Directions in Psychological Science, 21*, 215–220

Donkin, C., Brown, S., & Heathcote, A. (2011). Drawing conclusions from choice response time models: A tutorial using the linear ballistic accumulator. *Journal of Mathematical Psychology, 55*, 140–151.

Eidels, A., Donkin, C., Brown, S. D., & Heathcote, A. (2010). Converging measures of workload capacity. *Psychonomic Bulletin & Review, 17*, 763–771

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., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., & Breneiser, J. (2005). Multiple processes in prospective memory retrieval: factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology: General, 134*, 327-342.

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E. J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual review of psychology, 67*, 641-666.

Grunbaum, T., Oren, F., & Kyllingsbaek, S., (2021). A new cognitive model of long-term memory for intentions. *Cognition, 215,* 104817.

Guynn, M. J. (2003). A two-process model of strategic monitoring in event-based prospective memory: Activation/retrieval mode and checking. *International Journal of Psychology, 38*(4), 245-256.

Heathcote, A., Brown, S. D., & Wagenmakers, E. J. (2015a*).* An introduction to good practices in cognitive modeling. *In An introduction to model-based cognitive neuroscience (pp. 25-48).* Springer, New York, NY.

Heathcote, A., Loft, S., & Remington, R. W. (2015b). Slow down and remember to remember! A delay theory of prospective memory costs. *Psychological Review, 122*, 376-410.

Heathcote, A., Lin, Y. S., Reynolds, A., Strickland, L., Gretton, M., & Matzke, D. (2019). Dynamic models of choice. *Behavior research methods, 51*, 961-985.

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

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall, Inc

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*, 627–633.

Kuhlmann, B. G., & Rummel, J. (2014). Context-specific prospective memory processing: Evidence for flexible attention allocation adjustments after intention encoding. *Memory & Cognition*, *42*, 943-949

Kominsky, T. K., & Reese-Melancon, C. (2017). Effects of context expectation on prospective memory performance among older and younger adults. *Memory*, *25*(1), 122-131.

Kliegel, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2004). Importance effects on performance in event-based prospective memory tasks. *Memory, 12*, 553–561.

Kuhlmann, B. G., & Rummel, J. (2014). Context-specific prospective-memory processing: Evidence for flexible attention allocation adjustments after intention encoding. *Memory & Cognition*, *42*, 943-949.

Lerche, V., & Voss, A. (2020). When accuracy rates and mean response times lead to false conclusions: A simulation study based on the diffusion model. *The Quantitative Methods for Psychology, 16,* 107-119.

Lenth, R. V. (2021) emmeans: estimated marginal means, aka least-squares means (R package emmeans version 1.6.0). Comprehensive R Archive Network (CRAN).

Loft, S. (2014). Applying psychological science to examine prospective memory in simulated air traffic control. *Current Directions in Psychological Science, 23*, 326-331.

Loft, S., Dismukes, R. K., & Grundgeiger, T. (2019). Prospective memory in safety-critical work contexts. In Rummel, Jan (Ed) & McDaniel. Mark (Ed). *Current Issues in Memory: Prospective Memory.* (pp. 170-185). Taylor and Francis.

Loft, S., Finnerty, D., & Remington, R. W. (2011). Using spatial context to support prospective memory in simulated air traffic control. *Human Factors, 53*, 662-671.

Loft, S., & Remington, R. W. (2010). Prospective memory and task interference in a continuous monitoring dynamic display task. *Journal of Experimental Psychology: Applied, 16*, 145-167.

Loft, S., & Remington, R. W. (2013). Wait a second: Brief delays in responding reduce focality effects in event-based prospective memory. *Quarterly Journal of Experimental Psychology, 66*, 1432-1447.

Logan, G. D., Van Zandt, T., Verbruggen, F., & Wagenmakers, E.-J. (2014). On the ability to inhibit thought and action: General and special theories of an act of control. *Psychological Review, 121*, 66–95.

Lourenço, J. S., & Maylor, E. A. (2014). Is it relevant? Influence of trial manipulations of prospective memory context on task interference. *Quarterly Journal of Experimental Psychology, 67*, 687-702.

Lourenço, J. S., White, K., & Maylor, E. A. (2013). Target context specification can reduce costs in nonfocal prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1757-1764.

Marsh, R., Hicks, J., Cook, G., Hansen, J., & 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.

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

McDaniel, M. A., LaMontagne, P., Beck, S. M., Scullin, M. K., & Braver, T. S. (2013). Dissociable neural routes to successful prospective memory. *Psychological Science, 24*, 1791–1800

McElreath, R. (2020) Statistical Rethinking: *A Bayesian Course with Examples in R and STAN* (CRC Press).

Meier, B., Zimmermann, T. D., & Perrig, W. J. (2006). Retrieval experience in prospective memory: Strategic monitoring and spontaneous retrieval. *Memory, 14,* 872-889.

Miletić, S., Boag, R. J., & Forstmann, B. U. (2020). Mutual benefits: Combining reinforcement learning with sequential sampling models. *Neuropsychologia, 136*, 107261.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167–202

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods in Psychology, 4,* 61–64.

Navarro, D. J. (2019). Between the devil and the deep blue sea: Tensions between scientific judgement and statistical model selection. *Computational Brain & Behavior, 2,* 28-34.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review, 86*, 214–255

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104*, 266–300

Norman, D. (1981). Categorization of action slips. *Psychological Review, 88*, 1–15

Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 358–377.

Rummel J., & McDaniel M.A. (Eds.). (2019). Prospective memory. Abingdon, Oxon ; New York, NY: Routledge.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging, 19*, 278–289.

Reason, J. (1990). *Human error.* Cambridge, England: Cambridge University Press.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*. Sharpsburg, PA: Psychology Software Incorporated.

Scullin, M. K., McDaniel, M. A., & Shelton, J. T. (2013). The Dynamic Multiprocess Framework: Evidence from prospective memory with contextual variability. *Cognitive Psychology, 67*, 55-71.

Sewell, D. K., Jach, H. K., Boag, R. J., & Van Heer, C. A. (2019). Combining error-driven models of associative learning with evidence accumulation models of decision-making. *Psychonomic Bulletin & Review, 26*, 868-893.

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

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64,* 583–639.

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. (2017). *Prospective memory in context*. In B. H. Ross (Ed.), The Psychology of Learning and Motivation. New York: Academic Press.

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

Smith, R. E., Hunt, R. R., & Murray, A. E. (2017). Prospective memory in context: Moving through a familiar space. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 189-204.

Smith, R. E. & Skinner, D. (2019). Prospective memory in context: Methods, findings, and future directions. In J. Rummel & M. McDaniel (Eds.), *Current issues in memory: Prospective memory.* Taylor and Francis: Psychology Press

Strickland, L. (2023). Prospective Memory Decision Control: A Computational Model of Context Effects on Prospective Memory. https://doi.org/10.17605/OSF.IO/QEBMP

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

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

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

Strickland, L., Loft, S., & Heathcote, A. (2020). Investigating the effects of ongoing-task bias on prospective memory. *Quarterly Journal of Experimental Psychology, 73*, 1495-1513.

Strickland, L., Loft, S., & Heathcote, A. (2019a). Evidence accumulation modeling of event-based prospective memory. In Rummel, Jan (Ed) & McDaniel. Mark (Ed). *Current Issues in Memory: Prospective Memory*. (pp. 78-94). Taylor and Francis.

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

Voss, A., Voss, J., & Klauer, K. C. (2010). Separating response execution bias from decision bias: Arguments for an additional parameter in Ratcliff's diffusion model. *British Journal of Mathematical and Statistical Psychology, 63*, 539 –555**.**